

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: November 15, 2005, 23:53:58 / Search time 31202 Seconds
(without alignments)

11774.484 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcgagataataacc.....tattaactctgcactgcr 7582

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenBank:*

1: gb_ba:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7489.6	98.8	7582	BD196245	BD196245 Endogenet
2	7489.6	98.8	7582	BD267487	BD267487 Endogenet
3	7489.6	98.8	7582	AX000966	AX000966 Sequence
4	7489.6	98.8	7582	AX027480	AX027480 Sequence
5	6448	85.0	10222	AY101582	AY101582 Homo sapi
6	6448	85.0	10222	AY101585	AY101585 Homo sapi
7	6448	85.0	56093	AX329572	AX329572 Sequence
8	6448	85.0	56093	HSAC000064	HSAC000064 Human BAC
9	6448	85.0	149194	AC007565	AC007565 Homo sapi
10	6448	85.0	10222	AY101583	AY101583 Homo sapi
11	6440	84.9	10222	AY101584	AY101584 Homo sapi
12	6436	84.9	10499	BD221808	BD221808 Nucleic s
13	6436	84.9	10499	AX007980	AX007980 Sequence
14	6346.4	83.7	10229	AY101586	AY101586 Pan trogl
15	6344.8	83.7	10229	AY101587	AY101587 Pan trogl
16	6344.8	83.7	184675	AC145964	AC145964 Pan trogl
17	6269.6	82.7	10230	AY101588	AY101588 Gorilla g
18	6269.6	82.7	10230	AY101589	AY101589 Gorilla g
19	6037.8	79.6	10122	AY101590	AY101590 Pongo pyg

20	6025.8	79.5	10124	9	AY101591	AY101591 Pongo pyg
21	5956.8	78.6	10246	9	AY101593	AY101593 Hylobates
22	5950.4	78.5	10248	9	AY101592	AY101592 Hylobates
23	5824.2	76.8	158033	9	AY1018926	AY1018926 Homo sapi
24	5492	72.4	105989	9	AC008121	AC008121 Homo sapi
25	5492	72.4	110000	2	AC092510	AC092510 Papio anu
26	4755.6	62.7	187321	2	AC092510	AC092510 Papio anu
27	4302.2	56.7	114621	2	AC005187	AC005187 Homo sapi
28	4295.8	56.7	46575	2	AC080036	AC080036 Homo sapi
29	4162	54.9	111140	2	AL592310	AL592310 Human DNA
30	3944.6	52.0	183680	9	AC098859	AC098859 Homo sapi
31	3779.8	49.9	73070	9	AL590143	AL590143 Human DNA
32	3731.6	49.2	173788	9	AC022555	AC022555 Homo sapi
33	3731.6	49.2	186540	9	AC090341	AC090341 Homo sapi
34	3731.6	49.2	187837	9	AC079065	AC079065 Homo sapi
35	3716.6	49.0	160883	2	AL158814	AL158814 Homo sapi
36	3537.2	46.7	137947	9	HS4514	HS4514 Human DNA
37	3442.8	45.4	153444	2	AC040948	AC040948 Homo sapi
38	3442.8	45.4	205035	2	CNS00005	AL049870 Human chr
39	3438.2	45.3	101846	9	HS19711	AL031390 Human DNA
40	3438.2	45.3	224629	2	AC139260	AC139260 Homo sapi
41	3420.6	45.1	68693	9	AC108746	AC108746 Homo sapi
42	3367.4	44.4	118312	9	AC055713	AC055713 Homo sapi
43	3367.4	44.4	197549	2	AC145423	AC145423 Homo sapi
44	3300.6	43.5	169250	9	AP002812	AP002812 Homo sapi
45	3290.2	43.4	134412	2	AP002377	AP002377 Homo sapi

ALIGNMENTS

RESULT 1	BD196245	7582 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD196245				
DEFINITION	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.				
ACCESSION	BD196245.1	GI:33006015			
VERSION	BD196245.1	GI:33006015			
KEYWORDS	JP 2002512530-A/11.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 7582)				
AUTHORS	Beseme, F., Blond, J.L., Bouton, O., Mandrand, B. and Mallet, F.				
TITLE	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders				
JOURNAL	Patent: JP 2002512530-A 11 23-APR-2002;				
COMMENT	BIO MERIEUX				
OS	Unidentified				
PN	JP 2002512530-A/11				
PD	23-APR-2002				
PR	06-SEP-1998 JP 1999508244				
PR	07-JUL-1997 FR 97/08815				
PI	FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD MANDRAND,				
PI	FRANCOIS MALLET				
PC	C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC				
Strandness:	Single;				
CC	Topology: Linear;				
CC	Endogenetic retroviral sequences, associated with autoimmune diseases				
CC	diseases				
CC	and/or with pregnancy disorders				
FH	Key				
FT	source				
FT	1..7582				
FEATURES	Location/Qualifiers				
source	1..7582				
	/organism="Unidentified".				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	98.8%; Score 7489.6; DB 6; Length 7582;				

Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 CAACAATCGGGAATATAACCCAGGCAATGAGCTGCGACAACAGACCCCTTTGGGTCC 60
DB      1 CAACAATCGGGAATATAACCCAGGCAATGAGCTGCGACAACAGACCCCTTTGGGTCC 60
QY      61 CTTCCCTTTGTATAGGGAGCTGTTTCACTGATTTTCACTATTAATCTTGAACCTGCA 120
DB      61 CTTCCCTTTGTATAGGGAGCTGTTTCACTGATTTTCACTATTAATCTTGAACCTGCA 120
QY      121 CTTCTTGATGCATGTTTCTTAAGGCTGAGCTGAGCTTGTGCTCAACCTGCAACCACTGC 180
DB      121 CTTCTTGATGCATGTTTCTTAAGGCTGAGCTGAGCTTGTGCTCAACCTGCAACCACTGC 180
QY      181 TGTGTTCCACACCGACCGACCTGCGCTGACTCCCATCTCTGAGATCTGCAAGGGTGC 240
DB      181 TGTGTTCCACACCGACCGACCTGCGCTGACTCCCATCTCTGAGATCTGCAAGGGTGC 240
QY      241 CGGTGCTCTCTGATCCAGGAGGCGCCCATTTGCGGCTCCCAATGGGCTAAAGGGCTGC 300
DB      241 CGGTGCTCTCTGATCCAGGAGGCGCCCATTTGCGGCTCCCAATGGGCTAAAGGGCTGC 300
QY      301 CATTGTCCTGACAGGCTAGTGCCTGGTGTGTTCTAATTGAGCTGAACACTANTCACT 360
DB      301 CATTGTCCTGACAGGCTAGTGCCTGGTGTGTTCTAATTGAGCTGAACACTANTCACT 360
QY      361 GGGTTCATGCTTCTCTTCTGTGACCCAGGGCTTCTAATAKAACTATTAACATTTACCACA 420
DB      361 GGGTTCATGCTTCTCTTCTGTGACCCAGGGCTTCTAATAKAACTATTAACATTTACCACA 420
QY      421 TGGCCCAAGTTCATTCCTTTGGAATCCGTTGAGGSSACAAGACTCCAGGCTGAAGATAC 480
DB      421 TGGCCCAAGTTCATTCCTTTGGAATCCGTTGAGGSSACAAGACTCCAGGCTGAAGATAC 480
QY      481 GAGCTTGGCCACCATCTTGAAGGGGCTGCTACCTCTTGAAGTGTTCACCAACCATC 540
DB      481 GAGCTTGGCCACCATCTTGAAGGGGCTGCTACCTCTTGAAGTGTTCACCAACCATC 540
QY      541 TTGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGCAACMSPAAGGACATCC 600
DB      541 TTGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGCAACMSPAAGGACATCC 600
QY      601 MAAGTATGGGAAACGTTCCCGGCAACAACAAAACGCGCTTAACAGTATTCGGAAT 660
DB      601 MAAGTATGGGAAACGTTCCCGGCAACAACAAAACGCGCTTAACAGTATTCGGAAT 660
QY      661 TGGGAMCAATTTGACCTCAAGACATAAGAAAGAAAGAACTTATATTCTTCTGAGTGC 720
DB      661 TGGGAMCAATTTGACCTCAAGACATAAGAAAGAAAGAACTTATATTCTTCTGAGTGC 720
QY      721 GCTTGGAATCTCTGAGGAAAGTATTAATTATACACATCTTACAGCTAGACTTCTTTG 780
DB      721 GCTTGGAATCTCTGAGGAAAGTATTAATTATACACATCTTACAGCTAGACTTCTTTG 780
QY      781 TAGAAAAGGCAATGAGTGAAGTCCATAGTCAAACTTTCTTTTCAATTAAGACAA 840
DB      781 TAGAAAAGGCAATGAGTGAAGTCCATAGTCAAACTTTCTTTTCAATTAAGACAA 840
QY      841 CTACCAATTAATGTAAGAGTATTAATGCTTACAGAGAGCTTCAAGATCTACTCC 900
DB      841 CTACCAATTAATGTAAGAGTATTAATGCTTACAGAGAGCTTCAAGATCTACTCC 900
QY      901 CTATCCAGCATCCCGACTCTCTTCCCAATTAAGAGACCCCTTCAACCCAAATGG 960
DB      901 CTATCCAGCATCCCGACTCTCTTCCCAATTAAGAGACCCCTTCAACCCAAATGG 960
QY      961 TCCAAAAGAGATGACAAAAGGGTAAACAGTGAACCAAGATGCCAATTTCCCAAT 1020
DB      961 TCCAAAAGAGATGACAAAAGGGTAAACAGTGAACCAAGATGCCAATTTCCCAAT 1020
QY      1021 TATGACCCCTCCCAAGAGTGGAGGAAGAAATTCGCGCCAGGCAAGTGCATGTGCT 1080
DB      1021 TATGACCCCTCCCAAGAGTGGAGGAAGAAATTCGCGCCAGGCAAGTGCATGTGCT 1080
```

```
DB      1081 TTTTCTCCAGACTTAAGCAAAATTAAGCACTTATAGTAAATTCAGATAATTCCT 1140
QY      1081 TTTTCTCCAGACTTAAGCAAAATTAAGCACTTATAGTAAATTCAGATAATTCCT 1140
DB      1141 GATGCTATATGATGTTTATACAGGGTATGAGCAATCTTTGATCTGACATGAGAGAT 1200
QY      1141 GATGCTATATGATGTTTATACAGGGTATGAGCAATCTTTGATCTGACATGAGAGAT 1200
DB      1201 ATATATGATCTGCTTAATCAACACTAATCCCAATGAGAGAGGCAACCTAATCTGC 1260
QY      1201 ATATATGATCTGCTTAATCAACACTAATCCCAATGAGAGAGGCAACCTAATCTGC 1260
DB      1261 AGCTGAGAGTGTGGGATCTCTGATATCTAGTCAAGTCAATGATAGATGACACA 1320
QY      1261 AGCTGAGAGTGTGGGATCTCTGATATCTAGTCAAGTCAATGATAGATGACACA 1320
DB      1321 GAAGGAAGANATGATTTCCCAAGGCGACAGACGAGTTCCTAAGTCCATCTG 1380
QY      1321 GAAGGAAGANATGATTTCCCAAGGCGACAGACGAGTTCCTAAGTCCATCTG 1380
DB      1381 GGGACACAGAAATCAGTAAATGAGGAGATTTGGTCTGACAGATTGGCTAATCTGTGTC 1440
QY      1381 GGGACACAGAAATCAGTAAATGAGGAGATTTGGTCTGACAGATTGGCTAATCTGTGTC 1440
DB      1441 TASAAGGATTAAGAAAACTAAGAAATCTAAGAAATTAATCAATGATGTCACACATA 1500
QY      1441 TASAAGGATTAAGAAAACTAAGAAATCTAAGAAATTAATCAATGATGTCACACATA 1500
DB      1501 ACAAGAGGGAAGGAAAGAAATCTTACTGCTCTTCTGAGAGACTAAGGAGCATTTGAG 1560
QY      1501 ACAAGAGGGAAGGAAAGAAATCTTACTGCTCTTCTGAGAGACTAAGGAGCATTTGAG 1560
DB      1561 GAAAGCTGCTCTCTGACCTGACCTGACTCTTGAAGGCACTAATCTTAAAGGTAAGTT 1620
QY      1561 GAAAGCTGCTCTCTGACCTGACCTGACTCTTGAAGGCACTAATCTTAAAGGTAAGTT 1620
DB      1621 TATCACTAGTCAAGTCAAGCACTTGAAGAAACCTTCAAAAGCTGCGTGAAGCCGGAG 1680
QY      1621 TATCACTAGTCAAGTCAAGCACTTGAAGAAACCTTCAAAAGCTGCGTGAAGCCGGAG 1680
DB      1681 CAATACTTAAGAAACCTTATGAACCTTGGCAACTTGGCAACTTGGCAACTTGGCAACT 1740
QY      1681 CAATACTTAAGAAACCTTATGAACCTTGGCAACTTGGCAACTTGGCAACTTGGCAACT 1740
DB      1741 GAGCAGGCGGAAACAGGAAACGAGGATTAAGAAAGGCAACGCTTATGATGATGACCT 1800
QY      1741 GAGCAGGCGGAAACAGGAAACGAGGATTAAGAAAGGCAACGCTTATGATGATGACCT 1800
DB      1801 CAGGCAAGGGAATTTGAGAGGCTCTGAAAAGGAAAGCTGGGCAATTTGAATGCTTAA 1860
QY      1801 CAGGCAAGGGAATTTGAGAGGCTCTGAAAAGGAAAGCTGGGCAATTTGAATGCTTAA 1860
DB      1861 TAGGGCTGCTTCAAGGCGGTCTAAGAGCACTTTAAAAAGATGTCCAAAGTAAG 1920
QY      1861 TAGGGCTGCTTCAAGGCGGTCTAAGAGCACTTTAAAAAGATGTCCAAAGTAAG 1920
DB      1921 TAAGCGGCGCTTCTGTCAGAGCCCTTATTAAGGAAATCACTGAAAGGCCACTGCCC 1980
QY      1921 TAAGCGGCGCTTCTGTCAGAGCCCTTATTAAGGAAATCACTGAAAGGCCACTGCCC 1980
DB      1981 CAGGGGACAAAGGCTTTTGAATCAAGAGCACTAACCAGATGATCCAGAGCAGAGCTG 2040
QY      1981 CAGGGGACAAAGGCTTTTGAATCAAGAGCACTAACCAGATGATCCAGAGCAGAGCTG 2040
DB      2041 AGGGTGTCTGGGCGAAGCGCATCCCATGCAATCCCTCAAGAGCCCTGGGATGCTCT 2100
QY      2041 AGGGTGTCTGGGCGAAGCGCATCCCATGCAATCCCTCAAGAGCCCTGGGATGCTCT 2100
DB      2101 GACCATTTAGGGCGAAGAGGTTGTCTCTGAGACCTGAGCGGTCTTCTTAAGTCACT 2160
QY      2101 GACCATTTAGGGCGAAGAGGTTGTCTCTGAGACCTGAGCGGTCTTCTTAAGTCACT 2160
DB      2101 GACCATTTAGGGCGAAGAGGTTGTCTCTGAGACCTGAGCGGTCTTCTTAAGTCACT 2160
```

QY 2161 CTTCTGTCGGGACAACTGTCTCCAGATCTGTCACTAATTCTAGGGGGTCCNTAAGC 2220
Db 2161 CTTCTGTCCCGGACAACTGTCTCTCAGATCTGTCACTAATTCTAGGGGGTCCNTAAGC 2220
QY 2221 GGCAGTCACTAGATTACTTTTCCCGCACTAAGTATGAACTGGGGAGCTTTATCTTT 2280
Db 2221 GGCAGTCACTAGATTACTTTTCCCGCACTAAGTATGAACTGGGGAGCTTTATCTTT 2280
QY 2281 TCACATGCTTTTCTAATTATGCTTGAAGCCCACTACCTGTTTGAAGAGACATTTCTA 2340
Db 2281 TCACATGCTTTTCTAATTATGCTTGAAGCCCACTACCTGTTTGAAGAGACATTTCTA 2340
QY 2341 GCAAAAGCAGGGCCCATTAATACCTGTAACATAGAGAGAGAACCCGTTGTTGCC 2400
Db 2341 GCAAAAGCAGGGCCCATTAATACCTGTAACATAGAGAGAGAACCCGTTGTTGCC 2400
QY 2401 CCTGCTTGAGGAAGGAATTAATCTGTAAGTCTGGGCAACAGAGCAATATGACGAGC 2460
Db 2401 CCTGCTTGAGGAAGGAATTAATCTGTAAGTCTGGGCAACAGAGCAATATGACGAGC 2460
QY 2461 CAAAGATGCCCCCTGTTCAAGTTAACTAAAGATCCACTTCCCTACCAAA 2520
Db 2461 CAAAGATGCCCCCTGTTCAAGTTAACTAAAGATCCACTTCCCTACCAAA 2520
QY 2521 GGCAGTACCCCTCAGACCCAAAGGCCCAAGAGATTCAAAAGATTTGAAGACTTAA 2580
Db 2521 GGCAGTACCCCTCAGACCCAAAGGCCCAAGAGATTCAAAAGATTTGAAGACTTAA 2580
QY 2581 AAGCCCAAGGCTTAAGTAAACATGCAATACCTCTGCAAGTAATCCGTAGTGAG 2640
Db 2581 AAGCCCAAGGCTTAAGTAAACATGCAATACCTCTGCAAGTAATCCGTAGTGAG 2640
QY 2641 GAGGACAGAAACCCAGTGAAGAGGAGTTAGTCAAGATTCAGAAATTAATCAATG 2700
Db 2641 GAGGACAGAAACCCAGTGAAGAGGAGTTAGTCAAGATTCAGAAATTAATCAATG 2700
QY 2701 AGGCGTGTCTTTTATATCCAGCTGTACTAGCCCTTAATCTGTGMYTTCCCAATAC 2760
Db 2701 AGGCGTGTCTTTTATATCCAGCTGTACTAGCCCTTAATCTGTGMYTTCCCAATAC 2760
QY 2761 CAGAGGAACAGAGTGTTTACASTCTGACCTTAAAGATGCTTCTTGACATCCTG 2820
Db 2761 CAGAGGAACAGAGTGTTTACASTCTGACCTTAAAGATGCTTCTTGACATCCTG 2820
QY 2821 TACATCTGACTCAATCTTGTGCTTGAAGTACTTCAAAACCCARCAATCAAC 2880
Db 2821 TACATCTGACTCAATCTTGTGCTTGAAGTACTTCAAAACCCARCAATCAAC 2880
QY 2881 TCACTGCACTRTTTTACCAGAGGTTCAAGGATAGCCCCATCTAATTTGGCCAGGAT 2940
Db 2881 TCACTGCACTRTTTTACCAGAGGTTCAAGGATAGCCCCATCTAATTTGGCCAGGAT 2940
QY 2941 TAGCCCAAGACTTGAACATCTTGAAGTCTGTGCTTGTGTAAGTGAAG 3000
Db 2941 TAGCCCAAGACTTGAACATCTTGAAGTCTGTGCTTGTGTAAGTGAAG 3000
QY 3001 ATTACTTTTTCGCGCTTACAGAACTGTGACATCAAGGACCCAGGCTTTMA 3060
Db 3001 ATTACTTTTTCGCGCTTACAGAACTGTGACATCAAGGACCCAGGCTTTMA 3060
QY 3061 ATTCTCTGCGTACTGTGCTACAMGTTTCCAAASABARGTCTCTGTCTACAC 3120
Db 3061 ATTCTCTGCGTACTGTGCTACAMGTTTCCAAASABARGTCTCTGTCTACAC 3120
QY 3121 AGGTAAATTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
Db 3121 AGGTAAATTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
QY 3181 TCAGACTTAATCTGCTTATCTCATCYCAAAACCTTAAGCACTAAGGRRTTCTTG 3240
Db 3181 TCAGACTTAATCTGCTTATCTCATCYCAAAACCTTAAGCACTAAGGRRTTCTTG 3240

QY 3241 GCTTAAYAGYTTCTGCCGAATATGATTTCCCAAGTGTGGCBAATATGCCAGYCATTA 3300
Db 3241 GCTTAAYAGYTTCTGCCGAATATGATTTCCCAAGTGTGGCBAATATGCCAGYCATTA 3300
QY 3301 WATACASTAATTAAGAACTCAGAAAGCCAAATCCCATTAATTAATTAATTAATTAAT 3360
Db 3301 WATACASTAATTAAGAACTCAGAAAGCCAAATCCCATTAATTAATTAATTAATTAAT 3360
QY 3361 GYRAAGTGGCTTCCAGGCCCTTAAGAAAGGCTTTAAACCAAGCCAGTGTAAAT 3420
Db 3361 GYRAAGTGGCTTCCAGGCCCTTAAGAAAGGCTTTAAACCAAGCCAGTGTAAAT 3420
QY 3421 TGCCACRGGGCAAGATTTTSTTATATATCAAGAAAAAAGABAAVAGCTTREGA 3480
Db 3421 TGCCACRGGGCAAGATTTTSTTATATATCAAGAAAAAAGABAAVAGCTTREGA 3480
QY 3481 GTCCTTACACAGTCCBAGGAGAGAGCTTGCAACCTATGCGRYACTGASTAAGAAAT 3540
Db 3481 GTCCTTACACAGTCCBAGGAGAGAGCTTGCAACCTATGCGRYACTGASTAAGAAAT 3540
QY 3541 GATGTAGTGCAAGGTTGCTTCAATGTTTAYGGTATGTTGCGCATGACATYKTA 3600
Db 3541 GATGTAGTGCAAGGTTGCTTCAATGTTTAYGGTATGTTGCGCATGACATYKTA 3600
QY 3601 GTATCTGAAGCACTTAAATTAATACAGGGAGAGATCTTATCTGTGGAACATCTCAAG 3660
Db 3601 GTATCTGAAGCACTTAAATTAATACAGGGAGAGATCTTATCTGTGGAACATCTCAAG 3660
QY 3661 GTGAAYRGACTACTCATCTGTAAGGAGACTTGTGCTGTCAAGCACTTTACTTAA 3720
Db 3661 GTGAAYRGACTACTCATCTGTAAGGAGACTTGTGCTGTCAAGCACTTTACTTAA 3720
QY 3721 TTTCAAGCTCTATTAATTTGAAGGCGAGTGTGCGACTGTGCACTTGTCACTTAA 3780
Db 3721 TTTCAAGCTCTATTAATTTGAAGGCGAGTGTGCGACTGTGCACTTGTCACTTAA 3780
QY 3781 CCAGYCNCAATTTCTTCAGACATGAAGAAAGATTAATTAATTAATTAATTAATTAAT 3840
Db 3781 CCAGYCNCAATTTCTTCAGACATGAAGAAAGATTAATTAATTAATTAATTAATTAAT 3840
QY 3841 TCTCAAACTTAATGCCACTGAGGGGACTTGTAGAGTTCCTTGAATCCYGAATCC 3900
Db 3841 TCTCAAACTTAATGCCACTGAGGGGACTTGTAGAGTTCCTTGAATCCYGAATCC 3900
QY 3901 CAACCTGTATCTAGTGAAGTTCCTTGTGAAGAAAGCACTTGAAGGAGGATGTC 3960
Db 3901 CAACCTGTATCTAGTGAAGTTCCTTGTGAAGAAAGCACTTGAAGGAGGATGTC 3960
QY 3961 AGTGTCACTGATTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
Db 3961 AGTGTCACTGATTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
QY 4021 GCTRCAGAACTTAATAGCCCTCAATKGGGCACTAAGATTAAGAGAAAGGATTA 4080
Db 4021 GCTRCAGAACTTAATAGCCCTCAATKGGGCACTAAGATTAAGAGAAAGGATTA 4080
QY 4081 TATATATACAGCTTATATATGCTTAATCTTATCTTATGCTTATGCTTATGCTTAT 4140
Db 4081 TATATATACAGCTTATATATGCTTATATCTTATCTTATGCTTATGCTTATGCTTAT 4140
QY 4141 AGAAAGGAAATTCCTAATCTTGAAGGAACTATCAACATCGAGAAAGGCAATTAAG 4200
Db 4141 AGAAAGGAAATTCCTAATCTTGAAGGAACTATCAACATCGAGAAAGGCAATTAAG 4200
QY 4201 ATTATTAATGCGWGTACAGAAACCTTAAGAGTGAAGTCTTCACTGCGGGCTATCA 4260
Db 4201 ATTATTAATGCGWGTACAGAAACCTTAAGAGTGAAGTCTTCACTGCGGGCTATCA 4260
QY 4261 NAAAGGAAAGGAAAGGAAATTAAGGAAATGCTTATTAATTAATTAATTAATTAATTAAT 4320
Db 4261 NAAAGGAAAGGAAAGGAAATTAAGGAAATGCTTATTAATTAATTAATTAATTAATTAAT 4320
QY 4321 TGCAAGGAGGACCTCCTCAATTAAGAAATGCTTATTAATTAATTAATTAATTAATTAAT 4380

Db 4321 TGCAGGCGAGACCTCCATTTGAAATGCTTAATTAATCTCCCTTAGTATAGGTATATCC 4380
Qy 4381 CTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAAGATGGGAACTCAGAGG 4440
Db 4381 CTTCCGGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAAGATGGGAACTCAGAGG 4440
Qy 4441 CAGTTTTCTCCCTGGGGACGGTTAGCCACTGAGAGGAGGAAATATCTTTGCTGGCAAC 4500
Db 4441 CAGTTTTCTCCCTGGGGACGGTTAGCCACTGAGAGGAGGAAATATCTTTGCTGGCAAC 4500
Qy 4501 TATCCAAATGGAATTAATTAACCTTCATCAACCTTTCACTTAGCATGATAGCAC 4560
Db 4501 TATCCAAATGGAATTAATTAACCTTCATCAACCTTTCACTTAGCATGATAGCAC 4560
Qy 4561 CCATCATATGGCCAAATCATTAATTTATCTGACCAAGGCTTTTCAAAATCATCAAGCAT 4620
Db 4561 CCATCATATGGCCAAATCATTAATTTATCTGACCAAGGCTTTTCAAAATCATCAAGCAT 4620
Qy 4621 AKTCAGGGCTGTGAATGTGCGARABAAATATCCCTGCTTATGCGCAAGCTCTTC 4680
Db 4621 AKTCAGGGCTGTGAATGTGCGARABAAATATCCCTGCTTATGCGCAAGCTCTTC 4680
Qy 4681 AGGABAAACAAABAACAGGCCATTAACCTGABABAACTGGCAACTGATTTTACCAAG 4740
Db 4681 AGGABAAACAAABAACAGGCCATTAACCTGABABAACTGGCAACTGATTTTACCAAG 4740
Qy 4741 CCCAAACCTCAGGATTTCAATCTACTAGTCTGGGTABATCTTTCAGGGGTGGGCA 4800
Db 4741 CCCAAACCTCAGGATTTCAATCTACTAGTCTGGGTABATCTTTCAGGGGTGGGCA 4800
Qy 4801 RAAGGCTTCCCTGTATAGACAAABAAAGGCCAAAGGTATTAAGGCACTAATTTACGAA 4860
Db 4801 RAAGGCTTCCCTGTATAGACAAABAAAGGCCAAAGGTATTAAGGCACTAATTTACGAA 4860
Qy 4861 ATAAATCCAGATTTGSACTTCCCGAGGCTTACAGATGACAAATAGCCCTCTTCCAG 4920
Db 4861 ATAAATCCAGATTTGSACTTCCCGAGGCTTACAGATGACAAATAGCCCTCTTCCAG 4920
Qy 4921 GCCACAGTAAACCCAGGAGTATCCAGGCTTATAGGTATACATCTTACCTGCGCC 4980
Db 4921 GCCACAGTAAACCCAGGAGTATCCAGGCTTATAGGTATACATCTTACCTGCGCC 4980
Qy 4981 TGAAGGCCACAGTCTCAGGAGAGGTGAGAAATGAATGAATGAATGAATGAATGAATGA 5040
Db 4981 TGAAGGCCACAGTCTCAGGAGAGGTGAGAAATGAATGAATGAATGAATGAATGAATGA 5040
Qy 5041 AAAAAGCAAAACCCAGGAAACCCACCTCAGATGAGCTGTCTGTGCTATAGCTTAA 5100
Db 5041 AAAAAGCAAAACCCAGGAAACCCACCTCAGATGAGCTGTCTGTGCTATAGCTTAA 5100
Qy 5101 AGAATCTGCAACTTTTCCCAAAAAGAGAGCTTACGCTATAGCAAAATGCTGTATGAGAG 5160
Db 5101 AGAATCTGCAACTTTTCCCAAAAAGAGAGCTTACGCTATAGCAAAATGCTGTATGAGAG 5160
Qy 5161 CCCTTCAATACCAAGACCTTGTGCTGAGCCCAAGACGCCAATTAAGTTGAGACATTA 5220
Db 5161 CCCTTCAATACCAAGACCTTGTGCTGAGCCCAAGACGCCAATTAAGTTGAGACATTA 5220
Qy 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAGAAACCTATCCCTGAGAG 5280
Db 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAGAAACCTATCCCTGAGAG 5280
Qy 5281 AGGAAAAAGAACTATTTCCACCCWGTGACATGTATTAATCAATGCTCTTCTCTAAT 5340
Db 5281 AGGAAAAAGAACTATTTCCACCCWGTGACATGTATTAATCAATGCTCTTCTCTAAT 5340
Qy 5341 CCCATCCTCTAGATCATCTCTGGAGAGACCTTACCACTGATTTTATTAATTAATTAATTA 5400
Db 5341 CCCATCCTCTAGATCATCTCTGGAGAGACCTTACCACTGATTTTATTAATTAATTAATTA 5400
Qy 5401 CGGTTAAAGTGCTGAGTGAAGTCTTGAGATCATCACTTGAATCAATCTGAGATAC 5460

Db 5401 CGGTTAAAGTGCTGAGTGAAGTCTTGAGATCATCACTTGAATCAATCTGAGATAC 5460
Qy 5461 TGCCAAAGAAACCTGAAAATCCAGAGACCAACGCTATATCCCTGAGACCTTGAAGG 5520
Db 5461 TGCCAAAGAAACCTGAAAATCCAGAGACCAACGCTATATCCCTGAGACCTTGAAGG 5520
Qy 5521 ATTTGCGCTGCTCTTCAACAACAACAGAGAGGAAAGTAACTAAATCATTAATCCCCC 5580
Db 5521 ATTTGCGCTGCTCTTCAACAACAACAGAGAGGAAAGTAACTAAATCATTAATCCCCC 5580
Qy 5581 ATGSGCTCCCTTATCATATTTTCTCTKXASGTSTTTTAACTCTSTTTCACCTCACT 5640
Db 5581 ATGSGCTCCCTTATCATATTTTCTCTKXASGTSTTTTAACTCTSTTTCACCTCACT 5640
Qy 5641 GCACCCCCTCAATGCGCTGTATAGCAAGTATAGCTCCCTTACCMAGATTTATAGAGA 5700
Db 5641 GCACCCCCTCAATGCGCTGTATAGCAAGTATAGCTCCCTTACCMAGATTTATAGAGA 5700
Qy 5701 ATGCAAGCTCCCGAAATATTGAATGCCCATGTATAGAGTCTTSTTAAGGAAACCCC 5760
Db 5701 ATGCAAGCTCCCGAAATATTGAATGCCCATGTATAGAGTCTTSTTAAGGAAACCCC 5760
Qy 5761 ACCTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTTGCACCTTCTGATG 5820
Db 5761 ACCTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTTGCACCTTCTGATG 5820
Qy 5821 CATGCAATTAATCATTAATTTGAGACAGAAABAAATGAATTAATCTAGTTGCTGAGAGACT 5880
Db 5821 CATGCAATTAATCATTAATTTGAGACAGAAABAAATGAATTAATCTAGTTGCTGAGAGACT 5880
Qy 5881 GAGTCACTGTCTGTGAGCTTACTTCAACCAACTGGTATATGTGATGAGGGGTGAGTT 5940
Db 5881 GAGTCACTGTCTGTGAGCTTACTTCAACCAACTGGTATATGTGATGAGGGGTGAGTT 5940
Qy 5941 CAAGATCAGGCAAGAAABAAATGTAAAGAAATATCTCCAACTCAACGGGTATCAT 6000
Db 5941 CAAGATCAGGCAAGAAABAAATGTAAAGAAATATCTCCAACTCAACGGGTATCAT 6000
Qy 6001 GGCACTTACCCCTTACAAAGAGACTGATCTTCAAAACATACAGAAACCTCCGTACC 6060
Db 6001 GGCACTTACCCCTTACAAAGAGACTGATCTTCAAAACATACAGAAACCTCCGTACC 6060
Qy 6061 CATACTGCTGTGAGGCTTATTAATACCAACCTCAGTGGGCTTCATGAGGCTCTGGCC 6120
Db 6061 CATACTGCTGTGAGGCTTATTAATACCAACCTCAGTGGGCTTCATGAGGCTCTGGCC 6120
Qy 6121 CAAAACCTTACTAACTGTTGAGATGCTCTCCCTGAACTTCAGGCCATATGTTCAATC 6180
Db 6121 CAAAACCTTACTAACTGTTGAGATGCTCTCCCTGAACTTCAGGCCATATGTTCAATC 6180
Qy 6181 CCTGTACTGAAACATGGAACAATTCAGACACAGAAABAAACACCACTTCCGTTTATGA 6240
Db 6181 CCTGTACTGAAACATGGAACAATTCAGACACAGAAABAAACACCACTTCCGTTTATGA 6240
Qy 6241 GGAACCTCTGTTTCCAAATSTGAAATTAACCAATACCTCAAACTCAGCTGTATTAATTT 6300
Db 6241 GGAACCTCTGTTTCCAAATSTGAAATTAACCAATACCTCAAACTCAGCTGTATTAATTT 6300
Qy 6301 AGCAATACTACATACACAACCACTCCCAATGATCAGTGGGTAACTCTCCACACA 6360
Db 6301 AGCAATACTACATACACAACCACTCCCAATGATCAGTGGGTAACTCTCCACACA 6360
Qy 6361 ATAGCTGCTTACCTCAGGAAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6420
Db 6361 ATAGCTGCTTACCTCAGGAAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6420
Qy 6421 AATGCTCTTACAGATCTATGCTCTCTCATCTTATAGTCCCCCATGACATCTAC 6480
Db 6421 AATGCTCTTACAGATCTATGCTCTCTCATCTTATAGTCCCCCATGACATCTAC 6480
Qy 6481 ACTGAAACAAGATTATACAGTTATGTATCTAAGCCCGCAACAAGAGTACCATT 6540
Db 6481 ACTGAAACAAGATTATACAGTTATGTATCTAAGCCCGCAACAAGAGTACCATT 6540

Db 2581 AAGCCAAAGCTTAGTAAAAACATGCACTACTCCCTGCAAGTAAATTCCTGATGATGAG 2640
Qy 2641 GAGGACAGAAACCAGTGAAGTGAAGGTTAGTGAAGATCTCAGATTTATCAATGG 2700
Db 2641 GAGGACAGAAACCAGTGAAGTGAAGGTTAGTGAAGATCTCAGATTTATCAATGG 2700
Qy 2701 AGGCGTTGTCCTTTTATACCAGCTGACTAGCCCTTATACGTGMYTTCGAAATAC 2760
Db 2701 AGGCGTTGTCCTTTTATACCAGCTGACTAGCCCTTATACGTGMYTTCGAAATAC 2760
Qy 2761 CAGAGGAAGAGATGGTTTACSTCCGAGACCTTMAAGATGCTTCTTCGATCCCTG 2820
Db 2761 CAGAGGAAGAGATGGTTTACSTCCGAGACCTTMAAGATGCTTCTTCGATCCCTG 2820
Qy 2821 TACATCTGACTCTCAATCTTGTGTGCTTGAAGATCTTCAAAACCAACATCTCAAC 2880
Db 2821 TACATCTGACTCTCAATCTTGTGTGCTTGAAGATCTTCAAAACCAACATCTCAAC 2880
Qy 2881 TCACCTGACTRTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTGGCCAGGCAT 2940
Db 2881 TCACCTGACTRTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTGGCCAGGCAT 2940
Qy 2941 TAGCCCAAGACTGAGCAATGTCATACCTGAGACCTTGTCTTCCTTCGTAKGTGATG 3000
Db 2941 TAGCCCAAGACTGAGCAATGTCATACCTGAGACCTTGTCTTCCTTCGTAKGTGATG 3000
Qy 3001 ATTACTTTTGGCGCCRTTCAGAAACCTTGTGCTCAAGCCCAAGCCAGACCTCTTMA 3060
Db 3001 ATTACTTTTGGCGCCRTTCAGAAACCTTGTGCTCAAGCCCAAGCCAGACCTCTTMA 3060
Qy 3061 ATTTCTCGCYACCTGTGCTCAWGGTTTCCAAAGSABARBCCTCACTGTCTCAGAC 3120
Db 3061 ATTTCTCGCYACCTGTGCTCAWGGTTTCCAAAGSABARBCCTCACTGTCTCAGAC 3120
Qy 3121 AGGTTAAATCTTAGGCTAABATTAATCCAAAGGACACARGGCCCTCAGTGAAGAAATGA 3180
Db 3121 AGGTTAAATCTTAGGCTAABATTAATCCAAAGGACACARGGCCCTCAGTGAAGAAATGA 3180
Qy 3181 TCCAGCCTATACCTGCTTATCTCATCTCAATCCCTAAAGCACTAAGRRTTCTTGG 3240
Db 3181 TCCAGCCTATACCTGCTTATCTCATCTCAATCCCTAAAGCACTAAGRRTTCTTGG 3240
Qy 3241 GCTTAAVAGGYTTCGCGCAAAATGATTCGCCAGTGTGCGRAAATAGCCAGCYATTA 3300
Db 3241 GCTTAAVAGGYTTCGCGCAAAATGATTCGCCAGTGTGCGRAAATAGCCAGCYATTA 3300
Qy 3301 WATACASTAATTAAGGAATCTCAGAAAGCCATACCCATTATTAATGAATGAATGCTGA 3360
Db 3301 WATACASTAATTAAGGAATCTCAGAAAGCCATACCCATTATTAATGAATGAATGCTGA 3360
Qy 3361 GYWRBAAGTGGCTTCCAGGCCCCCTAAAGAGCCCTTAAACCAAGYCCAGTGTAAGYT 3420
Db 3361 GYWRBAAGTGGCTTCCAGGCCCCCTAAAGAGCCCTTAAACCAAGYCCAGTGTAAGYT 3420
Qy 3421 TGGCAACRGGGCAAGACTTTTSTTATATATCAAGAAAAAAAACAGAAVAGCTCRGGA 3480
Db 3421 TGGCAACRGGGCAAGACTTTTSTTATATATCAAGAAAAAAAACAGAAVAGCTCRGGA 3480
Qy 3481 GTTCCTTACACAGTCCRAAGGAGVAGCTTGCACCTRTGGCRATCTGASTAAGAAAT 3540
Db 3481 GTTCCTTACACAGTCCRAAGGAGVAGCTTGCACCTRTGGCRATCTGASTAAGAAAT 3540
Qy 3541 GATGTAGTGGCAAAAGGTTGCTTATTTTAAAGGAGTGTGCTGAGTGAAGCAATCTGAG 3600
Db 3541 GATGTAGTGGCAAAAGGTTGCTTATTTTAAAGGAGTGTGCTGAGTGAAGCAATCTGAG 3600
Qy 3601 GTATCTGAAGCACTTAAATATATACAGGGAAGATCTTACTGTGAGCATCTCAAGK 3660
Db 3601 GTATCTGAAGCACTTAAATATATACAGGGAAGATCTTACTGTGAGCATCTCAAGK 3660
Qy 3661 GTGAARGCATATCTACTGCTAAAGAGACTTGTGGCTGACAGCAACTTTTACTTAA 3720
Db 3661 GTGAARGCATATCTACTGCTAAAGAGACTTGTGGCTGACAGCAACTTTTACTTAA 3720

Qy 3721 TTTGAGGCTCTATTAATCTTAGAAGGCCAGTGTGCACTGTGCACTTGTGCAACTTTAAC 3780
Db 3721 TTTGAGGCTCTATTAATCTTAGAAGGCCAGTGTGCACTGTGCACTTGTGCAACTTTAAC 3780
Qy 3781 CCAGYCNATTTCTTCCAGACAAATGAAGAAAGATTAATTAATCTGCAACATTAAT 3840
Db 3781 CCAGYCNATTTCTTCCAGACAAATGAAGAAAGATTAATTAATCTGCAACATTAAT 3840
Qy 3841 TCTCAAACTATGCACTGAGGGACCTTTTATAGTTCCTTTGACTGATCCYACCTT 3900
Db 3841 TCTCAAACTATGCACTGAGGGACCTTTTATAGTTCCTTTGACTGATCCYACCTT 3900
Qy 3901 CAACTGTATACATGAGGAAGTTCTTTGTATGAAGAAAGCACTTGAAAGGGGTATGC 3960
Db 3901 CAACTGTATACATGAGGAAGTTCTTTGTATGAAGAAAGCACTTGAAAGGGGTATGC 3960
Qy 3961 AGTGTCACTGATTAATGAATATATTAATGAAATGCTTCACTCCAGAACTAGTCTTA 4020
Db 3961 AGTGTCACTGATTAATGAATATATTAATGAAATGCTTCACTCCAGAACTAGTCTTA 4020
Qy 4021 GCTRGCAAACTAATAGCCCTCAATGCGGCACTAAGATTTAGAGAAAGRAAAAGGAT 4080
Db 4021 GCTRGCAAACTAATAGCCCTCAATGCGGCACTAAGATTTAGAGAAAGRAAAAGGAT 4080
Qy 4081 TATATATACAGACTCTAATATGCTTAACCTAGTCCATGCTCAGTGMRCATATGSA 4140
Db 4081 TATATATACAGACTCTAATATGCTTAACCTAGTCCATGCTCAGTGMRCATATGSA 4140
Qy 4141 AGAAAGGAATTCCTAATCTTGAAGCAACCTATCAMAATCAGAAACCAATTAGAR 4200
Db 4141 AGAAAGGAATTCCTAATCTTGAAGCAACCTATCAMAATCAGAAACCAATTAGAR 4200
Qy 4201 ATTAATATYGGGCTGACAGAAACCTAAGAGTGAAGTCTTACATGCTGGGCTCA 4260
Db 4201 ATTAATATYGGGCTGACAGAAACCTAAGAGTGAAGTCTTACATGCTGGGCTCA 4260
Qy 4261 NAAAGGAAGRAAAGGAATASAAAGGAATGSCAAGCAKATTTGAAGCMAAAGAGC 4320
Db 4261 NAAAGGAAGRAAAGGAATASAAAGGAATGSCAAGCAKATTTGAAGCMAAAGAGC 4320
Qy 4321 TGCAGGCAAGACCTCTCATTAAGAAATGCTTAAATCTTCTTGTATAGGCTATCC 4380
Db 4321 TGCAGGCAAGACCTCTCATTAAGAAATGCTTAAATCTTCTTGTATAGGCTATCC 4380
Qy 4381 CTTCCGGGAAACCAAGCCCCCATGATCAGACAGAAAGAAAGATGCGGAACCTCAGAG 4440
Db 4381 CTTCCGGGAAACCAAGCCCCCATGATCAGACAGAAAGAAAGATGCGGAACCTCAGAG 4440
Qy 4441 CAGTTTCTCCCTCGGAGCGGTTAGCCACTGAAGAAAGGAAATACCTTTTGTGCTCAAC 4500
Db 4441 CAGTTTCTCCCTCGGAGCGGTTAGCCACTGAAGAAAGGAAATACCTTTTGTGCTCAAC 4500
Qy 4501 TATCCAAATGAAATTAATTAACCTTCAATCAACCTTCACTTGAAGCATGATGAC 4560
Db 4501 TATCCAAATGAAATTAATTAACCTTCAATCAACCTTCACTTGAAGCATGATGAC 4560
Qy 4561 CCATCARATGCGCAATCTATTTATCTGACCAAGCCCTTTCAAAACTATCAAGCARAT 4620
Db 4561 CCATCARATGCGCAATCTATTTATCTGACCAAGCCCTTTCAAAACTATCAAGCARAT 4620
Qy 4621 AKTCAGGCGCTGTGAATGTGCAABAAATATCCCTGCTATGCGCAAGCTCTTTC 4680
Db 4621 AKTCAGGCGCTGTGAATGTGCAABAAATATCCCTGCTATGCGCAAGCTCTTTC 4680
Qy 4681 AGABAAACAABAAACAAGCCATTAACCTGABAAABAATGCGAATGATTTTACCAAG 4740
Db 4681 AGABAAACAABAAACAAGCCATTAACCTGABAAABAATGCGAATGATTTTACCAAG 4740
Qy 4741 CCCAAACCTCAGAGATTTCAAGTATCTAAGTCTGGGTABATATCTTCAAGGTTGGCA 4800
Db 4741 CCCAAACCTCAGAGATTTCAAGTATCTAAGTCTGGGTABATATCTTCAAGGTTGGCA 4800

Db 6961 CTACTCTCTTGGACCTGTATCTTACCTCTTGTAACTTGTCTCTTCCAGATC 7020
 Qy 7021 GAAAGCTTAACTACAAATGAGAGCCCAAGATGAGTCAGACCTAAGATGTACCGAGA 7080
 Db 7021 GAAAGCTTAACTACAAATGAGAGCCCAAGATGAGTCAGACCTAAGATGTACCGAGA 7080
 Qy 7081 CCCCTGAGACGGGCTGTAGCCCAAGATGAGTCAGACCTAAGATGTACCGAGA 7140
 Db 7081 CCCCTGAGACGGGCTGTAGCCCAAGATGAGTCAGACCTAAGATGTACCGAGA 7140
 Qy 7141 GAGGAATCTCAGTGTACAGACCTCTCTAGAGCCCAATGAGAGAGAGATTAGAGC 7200
 Db 7141 GAGGAATCTCAGTGTACAGACCTCTCTAGAGCCCAATGAGAGAGAGATTAGAGC 7200
 Qy 7201 GGTGTGCGGCAACCTCCCAACAGACCTTAAAGTTTCTGTTGAGATGGGGGAGCTGAG 7260
 Db 7201 GGTGTGCGGCAACCTCCCAACAGACCTTAAAGTTTCTGTTGAGATGGGGGAGCTGAG 7260
 Qy 7261 GACAGACTAGCTGATTTCTTCTAGAGCTGATTAAGATCTTAAGCTTAAAGAGTGT 7320
 Db 7261 GACAGACTAGCTGATTTCTTCTAGAGCTGATTAAGATCTTAAGCTTAAAGAGTGT 7320
 Qy 7321 ACCACATCCACCTTTTAAACAGGGGCTTGCACCTTAACTACCAATCAGAGAG 7380
 Db 7321 ACCACATCCACCTTTTAAACAGGGGCTTGCACCTTAACTACCAATCAGAGAG 7380
 Qy 7381 CTCCTAAATGCTAATTAAGCAAGAGGTAAGAAATAGCCATCATTTATTCG 7440
 Db 7381 CTCCTAAATGCTAATTAAGCAAGAGGTAAGAAATAGCCATCATTTATTCG 7440
 Qy 7441 MTGAGAGCACAGGAGGAGCAATGATCGGATATTAACCAAGTTTGGAGCGGCA 7500
 Db 7441 MTGAGAGCACAGGAGGAGCAATGATCGGATATTAACCAAGTTTGGAGCGGCA 7500
 Qy 7501 CGGCAACCCCTTTGGGTCCCTTTGTATGGAGCTGTGTTTCAATGATTTTAC 7560
 Db 7501 CGGCAACCCCTTTGGGTCCCTTTGTATGGAGCTGTGTTTCAATGATTTTAC 7560
 Qy 7561 TCTATTAATCTTGACCTGCR 7582
 Db 7561 TCTATTAATCTTGACCTGCR 7582

Point

RESULT 3
 AX000966 7582 bp DNA 1 linear PAT 10-MAR-2000
 LOCUS AX000966
 DEFINITION Sequence 11 from Patent WO9902696.
 ACCESSION AX000966
 VERSION AX000966.1 GI:7241208
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 7582)
 AUTHORS Beseme, F. and Blond, J.
 TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
 DISEASES OR WITH PREGNANCY DISORDERS
 JOURNAL Patent: WO 9902696-A 11 21-JAN-1999;
 BIO MEDICAL (FR); BESEME, FREDERIC (FR)
 FEATURES
 source 1..7582
 /organism="unidentified"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3264"

Query Match 98.8%; Score 7489.6; DB 6; Length 7582;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAACATCGGATATTAACCCAGGACTTGGAGTGCACAGAGCCCTTTGGGTCC 60
 Db 1 CAACATCGGATATTAACCCAGGACTTGGAGTGCACAGAGCCCTTTGGGTCC 60

Qy 61 CTTCCTTTGATGAGAGCTTTTGAATGATATTTCACTATTAATCTTGAACCTGCA 120
 Db 61 CTTCCTTTGATGAGAGCTTTTGAATGATATTTCACTATTAATCTTGAACCTGCA 120
 Qy 121 CTCTTGTGTCATGTTTCTTAAGCTGAGCTGAGCTTGTGCTACCGTCCACCTG 180
 Db 121 CTCTTGTGTCATGTTTCTTAAGCTGAGCTGAGCTTGTGCTACCGTCCACCTG 180
 Qy 181 TGTTCGACACACCGGAACTGCTCCATCTCCCTGAGATCTGAGAGGTGTC 240
 Db 181 TGTTCGACACACCGGAACTGCTCCATCTCCCTGAGATCTGAGAGGTGTC 240
 Qy 241 CGCTGTCTCTGATGACAGAGAGCGCCATTCGCTCCCAATGGGCTTAAAGGCTTC 300
 Db 241 CGCTGTCTCTGATGACAGAGAGCGCCATTCGCTCCCAATGGGCTTAAAGGCTTC 300
 Qy 301 CATTGTCCTGACAGCTAAGTGCCTGGGTTTGTCTAATTTAGAGTGAACATTAACCT 360
 Db 301 CATTGTCCTGACAGCTAAGTGCCTGGGTTTGTCTAATTTAGAGTGAACATTAACCT 360
 Qy 361 GGGTTCATGTTCTCTTCTGACCAAGCTTCTTAATKAACCTTAACATTTACCA 420
 Db 361 GGGTTCATGTTCTCTTCTGACCAAGCTTCTTAATKAACCTTAACATTTACCA 420
 Qy 421 TGGCCCAAGATTCATTCTTGGAAATCCGAGAGGCAAGAACTCCAGGTCAGAGATAC 480
 Db 421 TGGCCCAAGATTCATTCTTGGAAATCCGAGAGGCAAGAACTCCAGGTCAGAGATAC 480
 Qy 481 GARGCTTGCACCATCTTGGAAAGCGGCTGTAACBCTTGTGAAGTGTTCACCACTATC 540
 Db 481 GARGCTTGCACCATCTTGGAAAGCGGCTGTAACBCTTGTGAAGTGTTCACCACTATC 540
 Qy 541 TTGGAGCTCTGTGACCAAGACCCCGGTTACATTTTGGCAACGAGGATTC 600
 Db 541 TTGGAGCTCTGTGACCAAGACCCCGGTTACATTTTGGCAACGAGGATTC 600
 Qy 601 MAAGTATGGAAGAAAGTTCCTCCGCAAGCAAAAAGCCCTTAAGAGATTCGGAAT 660
 Db 601 MAAGTATGGAAGAAAGTTCCTCCGCAAGCAAAAAGCCCTTAAGAGATTCGGAAT 660
 Qy 661 TGGAMCAATTTGACCTCAGACATTAAGAAAGAAAGCACTAATTTCTGCACTGCC 720
 Db 661 TGGAMCAATTTGACCTCAGACATTAAGAAAGAAAGCACTAATTTCTGCACTGCC 720
 Qy 721 GCTTGCACTCTGAGGAGAAATTAATTAACACACTTTAGACTGACTTCTTTTG 780
 Db 721 GCTTGCACTCTGAGGAGAAATTAATTAACACACTTTAGACTGACTTCTTTTG 780
 Qy 781 TAGAAAGGCAATGAGAGTGAAGTGCATTAAGTAACTTTCTTCAATTAAGAGCA 840
 Db 781 TAGAAAGGCAATGAGAGTGAAGTGCATTAAGTAACTTTCTTCAATTAAGAGCA 840
 Qy 841 CTCACAATTAATTAAGAAAGTGTATGCTTCAAGAGAGCTTCAAGTCTACCTCC 900
 Db 841 CTCACAATTAATTAAGAAAGTGTATGCTTCAAGAGAGCTTCAAGTCTACCTCC 900
 Qy 901 CTATCCAGCATCCCGGACTCTTCCCAATTAATTAAGAGAGCCCTTCAAGCAATG 960
 Db 901 CTATCCAGCATCCCGGACTCTTCCCAATTAATTAAGAGAGCCCTTCAAGCAATG 960
 Qy 961 TCCAAAGAGATGACAAAGAGGTAACAGGAAAGCAAGAGTGCATTAATTTCCCAAT 1020
 Db 961 TCCAAAGAGATGACAAAGAGGTAACAGGAAAGCAAGAGTGCATTAATTTCCCAAT 1020
 Qy 1021 TATGACCCCTCCCAAGAGTGGAGAGAGAAATTCGCTCCAGCAAGATGATGCTT 1080
 Db 1021 TATGACCCCTCCCAAGAGTGGAGAGAGAAATTCGCTCCAGCAAGATGATGCTT 1080
 Qy 1081 TTTTTCCTCCAGACTTAAAGAAATTAAGAAAGCACTTAAGTAAATTTCTCAATAAT 1140
 Db 1081 TTTTTCCTCCAGACTTAAAGAAATTAAGAAAGCACTTAAGTAAATTTCTCAATAAT 1140

QY 1141 GATGCTAATATGTGTTTACAAGGGTTAGACAATTCTTGTGATCTGACATGGAGAT 1200
| | | | |
Db 1141 GATGGCTAATATGTGTTTACAAGGGTTAGACAATTCTTGTGATCTGACATGGAGAT 1200
QY 1201 ATATATGTCATCTCTAATATGACACTAACCCCAATAGAGAGAGTCCACATACTGC 1260
| | | | |
Db 1201 ATATATGTCATCTCTAATATGACACTAACCCCAATAGAGAGAGTCCACATACTGC 1260
QY 1261 AGCTGTAGGTGTTGGCGATCTCTGATATCTCACTCAGGTCAATGGATATAGATGACACA 1320
| | | | |
Db 1261 AGCTGTAGGTGTTGGCGATCTCTGATATCTCACTCAGGTCAATGGATATAGATGACACA 1320
QY 1321 GAAGGAAGANNAATGATTTCCCAAGGCGACAGAGAGTTCACAGTTASACCTCATG 1380
| | | | |
Db 1321 GAAGGAAGANNAATGATTTCCCAAGGCGACAGAGAGTTCACAGTTASACCTCATG 1380
QY 1381 GGGACACAGAAATCAGTAAATGAGAGATGAGTCTGACAGACATTTGCTAATCTTGTC 1440
| | | | |
Db 1381 GGGACACAGAAATCAGTAAATGAGAGATGAGTCTGACAGACATTTGCTAATCTTGTC 1440
QY 1441 TASAAGACTAAGAAAACTASGAAGAAATCTAYGAATTACTCAATGATGTCACCA 1500
| | | | |
Db 1441 TASAAGACTAAGAAAACTASGAAGAAATCTAYGAATTACTCAATGATGTCACCA 1500
QY 1501 ACACAGGGGAAGGAGAAAAATCTACTGCTTTGAGAGAGACTAAGGAGGATTTAG 1560
| | | | |
Db 1501 ACACAGGGGAAGGAGAAAAATCTACTGCTTTGAGAGAGACTAAGGAGGATTTAG 1560
QY 1561 GAAGCGTCTCTCTGCTCAGCTGACCTTCTTGAAGGCACTAATCTTAAAGGTAAGT 1620
| | | | |
Db 1561 GAAGCGTCTCTCTGCTCAGCTGACCTTCTTGAAGGCACTAATCTTAAAGGTAAGT 1620
QY 1621 TATCACTCAGTCACTGACAGACATTAAGAAAACTTCAAAAGTCTGCGTAGCGCGAG 1680
| | | | |
Db 1621 TATCACTCAGTCACTGACAGACATTAAGAAAACTTCAAAAGTCTGCGTAGCGCGAG 1680
QY 1681 CAAAACTTAAAAACCCATTGAACCTTGGAACCTTGCGATTTTAAATGAAGATCAGGAG 1740
| | | | |
Db 1681 CAAAACTTAAAAACCCATTGAACCTTGGAACCTTGCGATTTTAAATGAAGATCAGGAG 1740
QY 1741 GACGAGCGGGAACGAGCAAAACGGGATTAAGAAAAAGGCAACGCTTTAGTCAATGACCT 1800
| | | | |
Db 1741 GACGAGCGGGAACGAGCAAAACGGGATTAAGAAAAAGGCAACGCTTTAGTCAATGACCT 1800
QY 1801 CAGGCAAGTGAATCTTTGAGGCTCTGAAAAAGGAAAAAGCTGGGCAATTTGAATGCTTAA 1860
| | | | |
Db 1801 CAGGCAAGTGAATCTTTGAGGCTCTGAAAAAGGAAAAAGCTGGGCAATTTGAATGCTTAA 1860
QY 1861 TAGGGCTTGTCTTCAAGTGGGCTCTCAAGAGACATTTAAAAAAGATTTGCCAATAGAG 1920
| | | | |
Db 1861 TAGGGCTTGTCTTCAAGTGGGCTCTCAAGAGACATTTAAAAAAGATTTGCCAATAGAG 1920
QY 1921 TAAAGCGCCCTTCTGCTCATGCCCCCTTATTTCAAGGAAATCACTGGAAGGCCACTGCC 1980
| | | | |
Db 1921 TAAAGCGCCCTTCTGCTCATGCCCCCTTATTTCAAGGAAATCACTGGAAGGCCACTGCC 1980
QY 1981 CAGGGGAACAAGGCTTTTGAAGTGAAGACCACTAACCAATGATGCCAGACAGAGACTG 2040
| | | | |
Db 1981 CAGGGGAACAAGGCTTTTGAAGTGAAGACCACTAACCAATGATGCCAGACAGAGACTG 2040
QY 2041 AGGGTGTCTGGGGGAACGCGCATCCATGCGCATCAACCTTCAAGAGCCCTGGGATCTT 2100
| | | | |
Db 2041 AGGGTGTCTGGGGGAACGCGCATCCATGCGCATCAACCTTCAAGAGCCCTGGGATCTT 2100
QY 2101 GACCATTAAGGGCCAGGAAGGTTGTCTCTGGAACATGATGCGGTCTTTAGTCTTACT 2160
| | | | |
Db 2101 GACCATTAAGGGCCAGGAAGGTTGTCTCTGGAACATGATGCGGTCTTTAGTCTTACT 2160
QY 2161 CTTCCTGCTCCGGAACAATGCTCTCAAGATCTGTCACTAATTTGAGGGGGTCTNTAAGACG 2220
| | | | |
Db 2161 CTTCCTGCTCCGGAACAATGCTCTCAAGATCTGTCACTAATTTGAGGGGGTCTNTAAGACG 2220
QY 2221 GGCAGTCACTAAGTACTTTTCCAGCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280
| | | | |

Db 2221 GGCAGTCACTAAGTACTTTTCCAGCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280
| | | | |
QY 2281 TCACATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTA 2340
| | | | |
Db 2281 TCACATGCTTTTCTAATTAATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTA 2340
QY 2341 GCAAAAGCAGGGGCCATTAATACACCGTGAACATAGAGAGAACACCCGTTTGTGTNC 2400
| | | | |
Db 2341 GCAAAAGCAGGGGCCATTAATACACCGTGAACATAGAGAGAACACCCGTTTGTGTNC 2400
QY 2401 CCTGTGAGAGAGAAATTAATCTGAAGTCTGGGCAAGAGACAAATAGAGAGAGC 2460
| | | | |
Db 2401 CCTGTGAGAGAGAAATTAATCTGAAGTCTGGGCAAGAGACAAATAGAGAGAGC 2460
QY 2461 CAAAGATGCCGTCCTGTTCAAGTTAACTAAGATTCACCTTCTTCCCTACCA 2520
| | | | |
Db 2461 CAAAGATGCCGTCCTGTTCAAGTTAACTAAGATTCACCTTCTTCCCTACCA 2520
QY 2521 GGCAGTACCCGCTCAGACCCCAAGGCCCAAGAGATTCACAAAGATTTAAGACCTTAA 2580
| | | | |
Db 2521 GGCAGTACCCGCTCAGACCCCAAGGCCCAAGAGATTCACAAAGATTTAAGACCTTAA 2580
QY 2581 AAGCCCAAGGCTTAAGTAAACCATGATGATTAATCCCTGACGTAAATCCGATGAGATTGAG 2640
| | | | |
Db 2581 AAGCCCAAGGCTTAAGTAAACCATGATGATTAATCCCTGACGTAAATCCGATGAGATTGAG 2640
QY 2641 GAAGCACAGAAACCAAGTGAACAGTGAAGGTTAGTGAAGATCTCAAGATTTCAATGG 2700
| | | | |
Db 2641 GAAGCACAGAAACCAAGTGAACAGTGAAGGTTAGTGAAGATCTCAAGATTTCAATGG 2700
QY 2701 AAGCGGTCTCTTTTAATACCAAGCTGTACTAAGCCCTAATCTGATTTCCCAATAC 2760
| | | | |
Db 2701 AAGCGGTCTCTTTTAATACCAAGCTGTACTAAGCCCTAATCTGATTTCCCAATAC 2760
QY 2761 CAGAGGAACAGAGTGTGTTTACASTCTGACCTTAAAGATGCTTCTGATCCCTG 2820
| | | | |
Db 2761 CAGAGGAACAGAGTGTGTTTACASTCTGACCTTAAAGATGCTTCTGATCCCTG 2820
QY 2821 TACATCTGACTCTCAATTTCTTGTGCTTGAAGATCTTCAACCCACATCTCAAC 2880
| | | | |
Db 2821 TACATCTGACTCTCAATTTCTTGTGCTTGAAGATCTTCAACCCACATCTCAAC 2880
QY 2881 TCACCTGGAATTTTAAACCCCAAGGTTCAAGGATGAGCCCATATTTGAGCAGGAT 2940
| | | | |
Db 2881 TCACCTGGAATTTTAAACCCCAAGGTTCAAGGATGAGCCCATATTTGAGCAGGAT 2940
QY 2941 TAGCCCAAGCTTGAGYCARITWTCACTACCTGGAACCTTGTCTTCTGRTAAGTGAAG 3000
| | | | |
Db 2941 TAGCCCAAGCTTGAGYCARITWTCACTACCTGGAACCTTGTCTTCTGRTAAGTGAAG 3000
QY 3001 ATTTACTTTTGGYGCCTTTTCAAAAACCTTGTGCAATCAAGCAACCAAGCTCTTMA 3060
| | | | |
Db 3001 ATTTACTTTTGGYGCCTTTTCAAAAACCTTGTGCAATCAAGCAACCAAGCTCTTMA 3060
QY 3061 ATTTCTTCCGACCTGCTGCAWGGTTTCAAAASBARAGTCACTGTGCTCAAGC 3120
| | | | |
Db 3061 ATTTCTTCCGACCTGCTGCAWGGTTTCAAAASBARAGTCACTGTGCTCAAGC 3120
QY 3121 AGGTAAATATCTTAGGCTTAATTAATCAAAAGCAACAGGCCCTCAAGTGAAGAAATTA 3180
| | | | |
Db 3121 AGGTAAATATCTTAGGCTTAATTAATCAAAAGCAACAGGCCCTCAAGTGAAGAAATTA 3180
QY 3181 TCAGGCTTACTGCTTATCTCATCTCAATCAAAACCTTAAAGCACTTAAGGRTTCTTG 3240
| | | | |
Db 3181 TCAGGCTTACTGCTTATCTCATCTCAATCAAAACCTTAAAGCACTTAAGGRTTCTTG 3240
QY 3241 GGTAAATAGGCTTCTGCGGAATATGATTTCCCAAGGTGAGCAATATAGCAGGCTTAA 3300
| | | | |
Db 3241 GGTAAATAGGCTTCTGCGGAATATGATTTCCCAAGGTGAGCAATATAGCAGGCTTAA 3300
QY 3301 WATACASTAATTAAGGAATCTGAAGGCAATACCATTTATTAAGATGAATGATGA 3360
| | | | |

Db 3301 WATACASTAATTAAAGAACTCAGAAAGCAATACCATTAAKTAAGATGAGYAMCTGAA 3360
Qy 3361 GYMBAAAGGCTTTCCAGAGCCCTTAAAGAGGCTTTAAACCAAGYCCCAAGTTAAAGT 3420
Db 3361 GYMBAAAGGCTTTCCAGAGCCCTTAAAGAGGCTTTAAACCAAGYCCCAAGTTAAAGT 3420
Qy 3421 TGCCAAACGGGCAAGACTTTTSTYATAYRTCAAGAAAAAAACAGBAAYAGCTCTRGGA 3480
Db 3421 TGCCAAACGGGCAAGACTTTTSTYATAYRTCAAGAAAAAAACAGBAAYAGCTCTRGGA 3480
Qy 3481 GTCCCTTACACAGRTCCBAGGGAAYAGCTTGCACCCYRTGGGRYACCTGASTPAAGAAAT 3540
Db 3481 GTCCCTTACACAGRTCCBAGGGAAYAGCTTGCACCCYRTGGGRYACCTGASTPAAGAAAT 3540
Qy 3541 GATGATGAGCAAGGGTTGACCTCATTTTAYGGGAGTGGTGGGCAAGTGCAGTYTA 3600
Db 3541 GATGATGAGCAAGGGTTGACCTCATTTTAYGGGAGTGGTGGGCAAGTGCAGTYTA 3600
Qy 3601 GATCTGAAGCAATTAAATATATACAGGGAAGATCTTACTGTGTGACATCTCATGAK 3660
Db 3601 GATCTGAAGCAATTAAATATATACAGGGAAGATCTTACTGTGTGACATCTCATGAK 3660
Qy 3661 GTGAAYGCAATCTCACTGCTAAAGAGACTTGTGGCTGCAACACGTGTTACTTAA 3720
Db 3661 GTGAAYGCAATCTCACTGCTAAAGAGACTTGTGGCTGCAACACGTGTTACTTAA 3720
Qy 3721 TATCAGGCTCTAATTACTTGAARGGCCAGTGTGCTGACCTTGTGCACTTAAAC 3780
Db 3721 TATCAGGCTCTAATTACTTGAARGGCCAGTGTGCTGACCTTGTGCACTTAAAC 3780
Qy 3781 CCAAGYCNATTTCTTCCAGACATGAAAGAAATAPAAAYTAACTGTCAACAAATAT 3840
Db 3781 CCAAGYCNATTTCTTCCAGACATGAAAGAAATAPAAAYTAACTGTCAACAAATAT 3840
Qy 3841 TCTCAAACTAATGCACTCGAGGGGACTTGTAGARGTTCCTTGAATCTGATCCGACTT 3900
Db 3841 TCTCAAACTAATGCACTCGAGGGGACTTGTAGARGTTCCTTGAATCTGATCCGACTT 3900
Qy 3901 CAACCTGTATACGATGGAAGTTCTTGTAGAAAAAGACTTGAAGYGGGGTATGC 3960
Db 3901 CAACCTGTATACGATGGAAGTTCTTGTAGAAAAAGACTTGAAGYGGGGTATGC 3960
Qy 3961 AGTGCTCATGATATGGAATATYTTGAAAGTATCCCTCATCTCAGAGAACTAGTCTTA 4020
Db 3961 AGTGCTCATGATATGGAATATYTTGAAAGTATCCCTCATCTCAGAGAACTAGTCTTA 4020
Qy 4021 GCTRGCAAGACTAATACCTCATYTKGGGCACTAGAAATYTAGAAGAAAGAAAGGTA 4080
Db 4021 GCTRGCAAGACTAATACCTCATYTKGGGCACTAGAAATYTAGAAGAAAGAAAGGTA 4080
Qy 4081 TATATATACAGACTCTAATATGCTTACCTAGTCTCATGCTGATGCAATGCAATATG 4140
Db 4081 TATATATACAGACTCTAATATGCTTACCTAGTCTCATGCTGATGCAATGCAATATG 4140
Qy 4141 AGAAAGGAATTCCTTAACTTCYAGAGAAACCTATCAMAATCAGGAAGCCATTAGAR 4200
Db 4141 AGAAAGGAATTCCTTAACTTCYAGAGAAACCTATCAMAATCAGGAAGCCATTAGAR 4200
Qy 4201 ATTAATATYGGGCTACAGAAACCTAAGAGTGGMACTTACCTGCTGAGGGTCTCA 4260
Db 4201 ATTAATATYGGGCTACAGAAACCTAAGAGTGGMACTTACCTGCTGAGGGTCTCA 4260
Qy 4261 NAAAGGAAGRAAAGGAATASAGRAAYTGCCAGCAKATYTTGAACMAAAGAGC 4320
Db 4261 NAAAGGAAGRAAAGGAATASAGRAAYTGCCAGCAKATYTTGAACMAAAGAGC 4320
Qy 4321 TGCAAGGCAAGACCTTCATTAAGATGCTTATTAACCTTCCTTATAGGGTAATCC 4380
Db 4321 TGCAAGGCAAGACCTTCATTAAGATGCTTATTAACCTTCCTTATAGGGTAATCC 4380
Qy 4381 TTCCCGGGAACCAAGCCCAAGTACTCAGCAGAGAAACAGAAATGGGGAACTTCACAG 4440
Db 4381 TTCCCGGGAACCAAGCCCAAGTACTCAGCAGAGAAACAGAAATGGGGAACTTCACAG 4440

Qy 4441 CAGTTTCTCCCTCGGGAGGTTAGGCACTGAAGAAAGGAAATACCTTTGCTGCAAC 4500
Db 4441 CAGTTTCTCCCTCGGGAGGTTAGGCACTGAAGAAAGGAAATACCTTTGCTGCAAC 4500
Qy 4501 TATCAATGGAATATTAATTAAACCTTCAATCAAACTTTCATTAAGCATGATAGAC 4560
Db 4501 TATCAATGGAATATTAATTAAACCTTCAATCAAACTTTCATTAAGCATGATAGAC 4560
Qy 4561 CCATCAATGCGCAATATTAATTAATGGAACAGGCTTTTCAAACTATCAAGCAT 4620
Db 4561 CCATCAATGCGCAATATTAATTAATGGAACAGGCTTTTCAAACTATCAAGCAT 4620
Qy 4621 AKTCAGGGCTGTGAATGTGCGCBAABAATATCCCTGCTAATGCGCAAGCTCTTC 4680
Db 4621 AKTCAGGGCTGTGAATGTGCGCBAABAATATCCCTGCTAATGCGCAAGCTCTTC 4680
Qy 4681 AGAARAAACAABAAACAGGCAATTAACCTGABARAACCTGCAACTGATTTTACCA 4740
Db 4681 AGAARAAACAABAAACAGGCAATTAACCTGABARAACCTGCAACTGATTTTACCA 4740
Qy 4741 CCCAACTCAGGGAATTTCAATTAATCTAATGCTGCTGCTGCTTCAAGGGTTGGCA 4800
Db 4741 CCCAACTCAGGGAATTTCAATTAATCTAATGCTGCTGCTGCTTCAAGGGTTGGCA 4800
Qy 4801 RAGGCTTCCCTGTAGGACAGAAAAAGCCCAAGAGTAATTAAGGCACTGATCATGA 4860
Db 4801 RAGGCTTCCCTGTAGGACAGAAAAAGCCCAAGAGTAATTAAGGCACTGATCATGA 4860
Qy 4861 ATTAATTCAGATTCGAGCTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG 4920
Db 4861 ATTAATTCAGATTCGAGCTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG 4920
Qy 4921 GCCAAGTAAACCAAGGAGATCCAGGGCTTATAGATATCACTTAACATGAGCC 4980
Db 4921 GCCAAGTAAACCAAGGAGATCCAGGGCTTATAGATATCACTTAACATGAGCC 4980
Qy 4981 TGAAGGCAAGCTCTCAGGGAAGTCAAGAAATGATGAAAYACTCAAGGACATCTA 5040
Db 4981 TGAAGGCAAGCTCTCAGGGAAGTCAAGAAATGATGAAAYACTCAAGGACATCTA 5040
Qy 5041 AAAAGGAACCCAGAGAAACCACTCAATGCTGCTGCTGCTTAAACCTTAA 5100
Db 5041 AAAAGGAACCCAGAGAAACCACTCAATGCTGCTGCTGCTTAAACCTTAA 5100
Qy 5101 AGAATCTGAACCTTCCCAAAAGAGAGACTTAAGCCCAATGCAAAAGCTGTATGAGG 5160
Db 5101 AGAATCTGAACCTTCCCAAAAGAGAGACTTAAGCCCAATGCAAAAGCTGTATGAGG 5160
Qy 5161 CCCTCATTAACCAATGACCTTGTGCTTGAACCAAGACAGCACTTATGTCAGACATCA 5220
Db 5161 CCCTCATTAACCAATGACCTTGTGCTTGAACCAAGACAGCACTTATGTCAGACATCA 5220
Qy 5221 CCTCTTAAACCAATATCAAAAGTTCTTAAACATTAAGGAAGAACTTATCCCTGAGAG 5280
Db 5221 CCTCTTAAACCAATATCAAAAGTTCTTAAACATTAAGGAAGAACTTATCCCTGAGAG 5280
Qy 5281 AGGGAAGAAAGATTAATCAACCCMWTGACATGATTAAGCAAGTCCCTTCYCTAAT 5340
Db 5281 AGGGAAGAAAGATTAATCAACCCMWTGACATGATTAAGCAAGTCCCTTCYCTAAT 5340
Qy 5341 CCCCATCCCTAAGTACATCTGGAAGAGACCTTACCTGATTTATTAATTAATTAATTA 5400
Db 5341 CCCCATCCCTAAGTACATCTGGAAGAGACCTTACCTGATTTATTAATTAATTAATTA 5400
Qy 5401 CGGTTAAAGTGGCTGGAAGTGTGATATCACTGATGATCAATCTGATAC 5460
Db 5401 CGGTTAAAGTGGCTGGAAGTGTGATATCACTGATGATCAATCTGATAC 5460
Qy 5461 TGCAAAAGGAACCTGAAATTCAGAGAGCAAGCTGATATTCCTGAGAG 5520
Db 5461 TGCAAAAGGAACCTGAAATTCAGAGAGCAAGCTGATATTCCTGAGAG 5520

QY 5521 ATTTGGGCTGCTCTTCAACACACAGAGGAAAGTAATCAATAATCCCC 5580
DB 5522 ATTTGGGCTGCTCTTCAACACACAGAGGAAAGTAATAATCAATAATCCCC 5580
QY 5581 ATGSGCTCCCTTATCATATTTTCTCTKTAATGTTSTTTTACCTTTTCACTCTACT 5640
DB 5581 ATGSGCTCCCTTATCATATTTTCTCTKTAATGTTSTTTTACCTTTTCACTCTACT 5640
QY 5641 GCACCCCTTCATGCGCTGATGACAGTAGTCTCCCTTACCMAGATTTCATGAGAGA 5700
DB 5641 GCACCCCTTCATGCGCTGATGACAGTAGTCTCCCTTACCMAGATTTCATGAGAGA 5700
QY 5701 ATGCAAGTCCCGGAAATTTTGAAGCCCATCGATAGAGTCTTTTAAAGGAGCCCC 5760
DB 5701 ATGCAAGTCCCGGAAATTTTGAAGCCCATCGATAGAGTCTTTTAAAGGAGCCCC 5760
QY 5761 ACCCTCACTGCCACACCCATATGCGCGCACTGCTATCACTGCGCACTTTGGATG 5820
DB 5761 ACCCTCACTGCCACACCCATATGCGCGCACTGCTATCACTGCGCACTTTGGATG 5820
QY 5821 CATGCAAAATCTCATTTATGAGACAGGAAAAATGATTAATCTAGTTTCTCGAGAGCTT 5880
DB 5821 CATGCAAAATCTCATTTATGAGACAGGAAAAATGATTAATCTAGTTTCTCGAGAGCTT 5880
QY 5881 GGAAGTCACTGTCTGTTGGAATTAATCACTTCACTGATGCTGATGAGGAGTGAAGT 5940
DB 5881 GGAAGTCACTGTCTGTTGGAATTAATCACTTCACTGATGCTGATGAGGAGTGAAGT 5940
QY 5941 CAAGATCAAGGCAAGAAAAATGTAAGAAATGTAATCTCCCACTCACGAGGATCAT 6000
DB 5941 CAAGATCAAGGCAAGAAAAATGTAAGAAATGTAATCTCCCACTCACGAGGATCAT 6000
QY 6001 GGCACCTCTGAGCCCTCAAAAGAGACTAGATCTCAAAACATCACTGAAACCTCCGATCC 6060
DB 6001 GGCACCTCTGAGCCCTCAAAAGAGACTAGATCTCAAAACATCACTGAAACCTCCGATCC 6060
QY 6061 CATACTGCTGCTGTAAGCTTATTAATPACACCTCACTGGGCTCATGAGGTCTGGCC 6120
DB 6061 CATACTGCTGCTGTAAGCTTATTAATPACACCTCACTGGGCTCATGAGGTCTGGCC 6120
QY 6121 CAATAACCTCTCACTGTAAGTATGCTCCCTGAACTTCAAGCCATATGTTCAATC 6180
DB 6121 CAATAACCTCTCACTGTAAGTATGCTCCCTGAACTTCAAGCCATATGTTCAATC 6180
QY 6181 CCTGTAACCTGAAACATGGAACCACTTCAGCAGAAATPAAACACACTTCCGTTTATGA 6240
DB 6181 CCTGTAACCTGAAACATGGAACCACTTCAGCAGAAATPAAACACACTTCCGTTTATGA 6240
QY 6241 GGAACCTTGTGTTCCAAATSTGGAATAACCATACCTCAACCTGCTGTAATAATT 6300
DB 6241 GGAACCTTGTGTTCCAAATSTGGAATAACCATACCTCAACCTGCTGTAATAATT 6300
QY 6301 AGCAATCTCATACATACACCAACCACTCCCAATGATCAGTGGATTAACCTCCCAACAA 6360
DB 6301 AGCAATCTCATACATACACCAACCACTCCCAATGATCAGTGGATTAACCTCCCAACAA 6360
QY 6361 ATAGTCTGCTACCTCAGAAATATTTTGTCTGTGTGTAACCTCAGCTTATGTTTGTG 6420
DB 6361 ATAGTCTGCTACCTCAGAAATATTTTGTCTGTGTGTAACCTCAGCTTATGTTTGTG 6420
QY 6421 AATGCTCTTCAAGATCTATGCTCTCTCATCTTCAATCTTAAAGCCCAATACCTCAAC 6480
DB 6421 AATGCTCTTCAAGATCTATGCTCTCTCATCTTCAATCTTAAAGCCCAATACCTCAAC 6480
QY 6481 ACTGAACAAGATTATACAGTTATGCTATCTTAAGCCCGCAACAAAGAGTACCAT 6540
DB 6481 ACTGAACAAGATTATACAGTTATGCTATCTTAAGCCCGCAACAAAGAGTACCAT 6540
QY 6541 CTTCTCTTTTGTATAGAGAGAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 6600
DB 6541 CTTCTCTTTTGTATAGAGAGAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 6600
QY 6601 ACAACCTCTACTCACTTCACTCAAACTATCTCAAGAACTTAATGAGGACATGGAACGG 6660

DB 6601 ACAACCTCTACTCACTTCACTCAAACTATCTCAAGAACTTAATGAGGACATGGAACGG 6660
QY 6661 GTGCGCACTCCCTGCTCACTTTGCAAGATCAACTTAATCTCTTACGACAGATGCTCT 6720
DB 6661 GTGCGCACTCCCTGCTCACTTTGCAAGATCAACTTAATCTCTTACGACAGATGCTCT 6720
QY 6721 CBAATCGAAGAGCTTTAAGCTGTAACGCGCTGABAGAGGAGGAACTTTTATTTTAA 6780
DB 6721 CBAATCGAAGAGCTTTAAGCTGTAACGCGCTGABAGAGGAGGAACTTTTATTTTAA 6780
QY 6781 GGGGAAGAAATGCTGTTATATGTTAATCAATCGGAATGCTCATGAGAAATTTAAAGAA 6840
DB 6781 GGGGAAGAAATGCTGTTATATGTTAATCAATCGGAATGCTCATGAGAAATTTAAAGAA 6840
QY 6841 ATTCSAGATCGAATTAACAGTAAAGAGAGAGCTTCGAAACACTGACCCCTGGGCTC 6900
DB 6841 ATTCSAGATCGAATTAACAGTAAAGAGAGAGCTTCGAAACACTGACCCCTGGGCTC 6900
QY 6901 CTCAGCCBATGGAATGCGCTGGAATCTCCCTTCTTGAACCTTACGAGCTATTAATTTG 6960
DB 6901 CTCAGCCBATGGAATGCGCTGGAATCTCCCTTCTTGAACCTTACGAGCTATTAATTTG 6960
QY 6961 CTACTCTCTTTGGAACCTGATCTTTTBACTTCTTGTATTAATCTTCTTCCAGATC 7020
DB 6961 CTACTCTCTTTGGAACCTGATCTTTTBACTTCTTGTATTAATCTTCTTCCAGATC 7020
QY 7021 GAACTGTBAATTAACAAATGAGCCCAAGATCAGTCCAAACTAAGATCTACCGAGA 7080
DB 7021 GAACTGTBAATTAACAAATGAGCCCAAGATCAGTCCAAACTAAGATCTACCGAGA 7080
QY 7081 CCCCTGAGCCGCTGTATGAGCCCAAGATCAGTCCAAACTAAGATCTACCGAGA 7140
DB 7081 CCCCTGAGCCGCTGTATGAGCCCAAGATCAGTCCAAACTAAGATCTACCGAGA 7140
QY 7141 GAGAAATCTCACTGCAACCACTTACTACCTGCCCCCAATCAAGCAGGAGCACTTAAGC 7200
DB 7141 GAGAAATCTCACTGCAACCACTTACTACCTGCCCCCAATCAAGCAGGAGCACTTAAGC 7200
QY 7201 GGTGTGAGGCAACCTTCCCAACAGCACTTAAGTTTCTGTTGAGATGGGGAGCTGAGA 7260
DB 7201 GGTGTGAGGCAACCTTCCCAACAGCACTTAAGTTTCTGTTGAGATGGGGAGCTGAGA 7260
QY 7261 GACAGACTGCTGAGATTTCTTAGGCTGATTAAGATCTTAACTGAGTGGGAAAGGTG 7320
DB 7261 GACAGACTGCTGAGATTTCTTAGGCTGATTAAGATCTTAACTGAGTGGGAAAGGTG 7320
QY 7321 ACCAATCCACTTTTAAACAGGGGCTTGGCACTTAGTCAACCTGACCAATCAGAGAG 7380
DB 7321 ACCAATCCACTTTTAAACAGGGGCTTGGCACTTAGTCAACCTGACCAATCAGAGAG 7380
QY 7381 CTCATTAATAATGCTAATTAGGCAAGACAGAGGTTAAAGAAATGCCAATCATTTATTCG 7440
DB 7381 CTCATTAATAATGCTAATTAGGCAAGACAGAGGTTAAAGAAATGCCAATCATTTATTCG 7440
QY 7441 MTGAGACACAGACAGAGGAGCAATGATCGGGAATATAACCAAGTGTTCAGCGGCA 7500
DB 7441 MTGAGACACAGACAGAGGAGCAATGATCGGGAATATAACCAAGTGTTCAGCGGCA 7500
QY 7501 CGGCAACCCCTTTGGGTCCCTCTCTTTGATGAGGCTGTGTTTCACTCAATTTTCAAC 7560
DB 7501 CGGCAACCCCTTTGGGTCCCTCTCTTTGATGAGGCTGTGTTTCACTCAATTTTCAAC 7560
QY 7561 TCTATTAATCTTGCARCTGCR 7582
DB 7561 TCTATTAATCTTGCARCTGCR 7582
RESULT 4
AX027480 7582 bp DNA linear PAT 16-Sep-2000
LOCUS AX027480
DEFINITION Sequence 30 from Patent FR2788784.
ACCESSION AX027480

VERSION AX027480.1 GI:10188444
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 Mallet, F., Voisard, C. and Paranhos, B.G.
AUTHORS Patent: FR 2788784-A 30 28-JUL-2000;
JOURNAL BIO MERIEUX (FR)
FEATURES Location/Qualifiers
source 1. 7582
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.8%; Score 7489.6; DB 6; Length 7582;
Best Local Similarity 100.0%; Pired. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACAATCGGGATATTAACCCAGGCAATTCAGCTGGGAGACAGACGCCCCCTTTGGGTCC 60
DB 1 CAACAATCGGGATATTAACCCAGGCAATTCAGCTGGGAGACAGACGCCCCCTTTGGGTCC 60
QY 61 CTTCCTTTGATGGAGCTGTTTCAATGCTATTTCACTCTATTAATCTTGGCACTGCA 120
DB 61 CTTCCTTTGATGGAGCTGTTTCAATGCTATTTCACTCTATTAATCTTGGCACTGCA 120
QY 121 CTCTTCTGTCATGTTTCTTAAGGCTGAGCTGAGCTTTTGGCTCACGCTCCACACTGC 180
DB 121 CTCTTCTGTCATGTTTCTTAAGGCTGAGCTGAGCTTTTGGCTCACGCTCCACACTGC 180
QY 122 CTCTTCTGTCATGTTTCTTAAGGCTGAGCTGAGCTTTTGGCTCACGCTCCACACTGC 180
DB 122 CTCTTCTGTCATGTTTCTTAAGGCTGAGCTGAGCTTTTGGCTCACGCTCCACACTGC 180
QY 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCCGAGACCTCCGAGGGGTGC 240
DB 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCCGAGACCTCCGAGGGGTGC 240
QY 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCCGAGACCTCCGAGGGGTGC 240
DB 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCCGAGACCTCCGAGGGGTGC 240
QY 241 CGCTGTCCTCTGATCCAGGAGCGCCCAATGGCCGCTCCCAATTTGGGCTAAAGGCTTC 300
DB 241 CGCTGTCCTCTGATCCAGGAGCGCCCAATGGCCGCTCCCAATTTGGGCTAAAGGCTTC 300
QY 241 CGCTGTCCTCTGATCCAGGAGCGCCCAATGGCCGCTCCCAATTTGGGCTAAAGGCTTC 300
DB 241 CGCTGTCCTCTGATCCAGGAGCGCCCAATGGCCGCTCCCAATTTGGGCTAAAGGCTTC 300
QY 301 CATGTGCTGCAACGCTAAGTGGCTGGTGTGTTCTAATTAAGCTGAACACTAATCACT 360
DB 301 CATGTGCTGCAACGCTAAGTGGCTGGTGTGTTCTAATTAAGCTGAACACTAATCACT 360
QY 361 GGGTTCCATGCTCTCTCTGTAACCCAGGCTTCTAATAAATACTAATACTTAACCA 420
DB 361 GGGTTCCATGCTCTCTCTGTAACCCAGGCTTCTAATAAATACTAATACTTAACCA 420
QY 421 TGGCCCAAGATTCATTCCTTGGAAATCGTAGGSCAAGAACTCCAGCTCAGAGAAATAC 480
DB 421 TGGCCCAAGATTCATTCCTTGGAAATCGTAGGSCAAGAACTCCAGCTCAGAGAAATAC 480
QY 481 GARGCTTGGCAACCATCTTGAAGCGGCTGCTACCTCTTGAAGTGGTTCACACCATTC 540
DB 481 GARGCTTGGCAACCATCTTGAAGCGGCTGCTACCTCTTGAAGTGGTTCACACCATTC 540
QY 541 TTGGGAGCTCTGAGCAAGAACCCCGGCTTACATTTTGGCRAACMSAAGCACTCC 600
DB 541 TTGGGAGCTCTGAGCAAGAACCCCGGCTTACATTTTGGCRAACMSAAGCACTCC 600
QY 601 MAAGTATGGGAAACGTTCCCGGCAAGACAAAGAGCCCTTAAGACGATTTCTGGAAT 660
DB 601 MAAGTATGGGAAACGTTCCCGGCAAGACAAAGAGCCCTTAAGACGATTTCTGGAAT 660
QY 661 TGGGAAATTTGACCCCTGAGACATTAAGAAAGAAAGCACTTAATTTCTGCAAGTCC 720
DB 661 TGGGAAATTTGACCCCTGAGACATTAAGAAAGAAAGCACTTAATTTCTGCAAGTCC 720
QY 721 GCTTGGCACTCTCGAGGAAAGTATTAATTAACCACTTTTACAGTGAATCTTTTG 780
DB 721 GCTTGGCACTCTCGAGGAAAGTATTAATTAACCACTTTTACAGTGAATCTTTTG 780

QY 781 TAGAAAAGCAATGAGTGAAGTGCATTAAGTAACTTTCTTTCAATTAAGACAA 840
DB 781 TAGAAAAGCAATGAGTGAAGTGCATTAAGTAACTTTCTTTCAATTAAGACAA 840
QY 841 CTCACAAATTAATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
DB 841 CTCACAAATTAATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
QY 901 CTATCCAGATCCCGGACCTCTTCCCAATTAATTAAGACCCCTTCAACCAATTCG 960
DB 901 CTATCCAGATCCCGGACCTCTTCCCAATTAATTAAGACCCCTTCAACCAATTCG 960
QY 961 TCCAAAAGAGATAGACAAAGGCTTAAACAGTAAACCAAGAGTGCATATTCCTCAAT 1020
DB 961 TCCAAAAGAGATAGACAAAGGCTTAAACAGTAAACCAAGAGTGCATATTCCTCAAT 1020
QY 1021 TATGACCCCTCCCAAGCAGTGGAGAGAAATTCGGCCAGCCAGATGTCATGTCCT 1080
DB 1021 TATGACCCCTCCCAAGCAGTGGAGAGAAATTCGGCCAGCCAGATGTCATGTCCT 1080
QY 1081 TTTTCTCCAGACTTAAGCAATTAAGCACTTAAGTAAATTCAGATTAATTCCT 1140
DB 1081 TTTTCTCCAGACTTAAGCAATTAAGCACTTAAGTAAATTCAGATTAATTCCT 1140
QY 1141 GATGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
DB 1141 GATGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
QY 1201 ATATATGTCATCTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
DB 1201 ATATATGTCATCTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
QY 1261 AGCTGAGGTTTGGGATCTGCTGATCTGATGATCTGATGATGATGATGATGATGAT 1320
DB 1261 AGCTGAGGTTTGGGATCTGCTGATCTGATGATCTGATGATGATGATGATGATGAT 1320
QY 1321 GAAGGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
DB 1321 GAAGGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
QY 1381 GGGACACAGAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
DB 1381 GGGACACAGAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
QY 1441 TASAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
DB 1441 TASAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
QY 1501 ACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 ACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 GAAGGCTGCTCTGCTGACCTGATCTTCTGAGGCACTAATCTTAAAGGTAAGTT 1620
DB 1561 GAAGGCTGCTCTGCTGACCTGATCTTCTGAGGCACTAATCTTAAAGGTAAGTT 1620
QY 1621 TATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1621 TATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 CAATACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
DB 1681 CAATACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
QY 1741 GAGCAGGCGGAGCAGGACAAAGGAGTTAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
DB 1741 GAGCAGGCGGAGCAGGACAAAGGAGTTAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
QY 1801 CAGGCAAGTGAATTTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
DB 1801 CAGGCAAGTGAATTTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
QY 1861 TAGGCTGCTCTTCCAGTGGGCTTCAAGGACCTTTAAAGAAATGTCACAGTGAAG 1920

QY 4081 TATATATACAGACTCTBARTATGCTVACCTAGTCNTCCATGCCCATGMRGCAATATGSAR 4140
DB 4081 TATATATACAGACTCTBARTATGCTVACCTAGTCNTCCATGCCCATGMRGCAATATGSAR 4140
QY 4141 AGAAAGGAAATTCCTAACTTCYAGRGAACAACCTATCMAACATCAGAGACCTTAGGAR 4200
DB 4141 AGAAAGGAAATTCCTAACTTCYAGRGAACAACCTATCMAACATCAGAGACCTTAGGAR 4200
QY 4201 ATATTATYTGCGWGTACAGAAACCTABAGAGTGMAGTCTTACACGTCGCGGTCTCA 4260
DB 4201 ATATTATYTGCGWGTACAGAAACCTABAGAGTGMAGTCTTACACGTCGCGGTCTCA 4260
QY 4261 NAAAGGAAAGRAAAGGAAATASAGRAAYTGCCAGACAKATTTAGACMAAAAGAGC 4320
DB 4261 NAAAGGAAAGRAAAGGAAATASAGRAAYTGCCAGACAKATTTAGACMAAAAGAGC 4320
QY 4321 TGCAGAGCAGAGCCCTCCATTAGAAATGCTTATTTAACTTCCCTTATATAGGTAATCC 4380
DB 4321 TGCAGAGCAGAGCCCTCCATTAGAAATGCTTATTTAACTTCCCTTATATAGGTAATCC 4380
QY 4381 CTTCCGGGAAACCAAGCCCAAGTACTCAGACAGAGAAACAGAAATGGGAACTCAGAGG 4440
DB 4381 CTTCCGGGAAACCAAGCCCAAGTACTCAGACAGAGAAACAGAAATGGGAACTCAGAGG 4440
QY 4441 CAGTTTCTCCCTCGGAGCGTTAGGCACTGAAGAAAGGAAATACTTTTGCTGCTGAC 4500
DB 4441 CAGTTTCTCCCTCGGAGCGTTAGGCACTGAAGAAAGGAAATACTTTTGCTGCTGAC 4500
QY 4501 TATCCAAATGAAATTAATTAAACCTTCATCAAACTTTCACTTAGGCATGATAGAC 4560
DB 4501 TATCCAAATGAAATTAATTAAACCTTCATCAAACTTTCACTTAGGCATGATAGAC 4560
QY 4561 CCATCBAATGSCCAATTAATTAATCTGACAGAGCCTTTCAAACTCAAGCBAAT 4620
DB 4561 CCATCBAATGSCCAATTAATTAATCTGACAGAGCCTTTCAAACTCAAGCBAAT 4620
QY 4621 AATCAGGCGCTGTGAATGTGACABARAATAATCCCTGCTATATGSCCAAGCTCTTC 4680
DB 4621 AATCAGGCGCTGTGAATGTGACABARAATAATCCCTGCTATATGSCCAAGCTCTTC 4680
QY 4681 AGGARAACAARAACAAGGCCATTACCTGABARAATCTGSCAACTGATTTTACCAAG 4740
DB 4681 AGGARAACAARAACAAGGCCATTACCTGABARAATCTGSCAACTGATTTTACCAAG 4740
QY 4741 CCCAAACCTCAGGAGATTTCACTATCTAATGCTGCTGATATATCTTCAAGGGTGGCA 4800
DB 4741 CCCAAACCTCAGGAGATTTCACTATCTAATGCTGCTGATATATCTTCAAGGGTGGCA 4800
QY 4801 RAGGCTTCCCTGTAGGAAGAAAGGCCAAGAGGTAAATAAGGCACTAGTTCAATGA 4860
DB 4801 RAGGCTTCCCTGTAGGAAGAAAGGCCAAGAGGTAAATAAGGCACTAGTTCAATGA 4860
QY 4861 ATATATCCAGATTCGGAATTCCTCCAGAGCTTACAGAGTGAACAATAGCCTGCTTCAG 4920
DB 4861 ATATATCCAGATTCGGAATTCCTCCAGAGCTTACAGAGTGAACAATAGCCTGCTTCAG 4920
QY 4921 GCCACAGTAACCAAGGAGTATCCCAAGGCTTATGATATACATCTTACCTGAGCC 4980
DB 4921 GCCACAGTAACCAAGGAGTATCCCAAGGCTTATGATATACATCTTACCTGAGCC 4980
QY 4981 TGAAGGCCACAGTCTCAGGGAAGTGCAGAAATGAATAAAYACTCAAGAGACATCTA 5040
DB 4981 TGAAGGCCACAGTCTCAGGGAAGTGCAGAAATGAATAAAYACTCAAGAGACATCTA 5040
QY 5041 AAAAGCAAAACCAAGAAACCACTCACAATGCGCTGTCTGTGCTATAGCCTTAA 5100
DB 5041 AAAAGCAAAACCAAGAAACCACTCACAATGCGCTGTCTGTGCTATAGCCTTAA 5100
QY 5101 AGAATCTGCACTTTCCCAAAAGCAGACTTAGCCCAATAGAAATGCTGTATGGAAG 5160
DB 5101 AGAATCTGCACTTTCCCAAAAGCAGACTTAGCCCAATAGAAATGCTGTATGGAAG 5160

QY 5161 CCCTTCATTAACCAATGACCTGTGTGACCCCAAGACAGCCAACTTAGTGCAGACATCA 5220
DB 5161 CCCTTCATTAACCAATGACCTGTGTGACCCCAAGACAGCCAACTTAGTGCAGACATCA 5220
QY 5221 CCTCTTACCAATATCAACAAGTTCTTAAACATTAACAAGACCTATCCTGAGAG 5280
DB 5221 CCTCTTACCAATATCAACAAGTTCTTAAACATTAACAAGACCTATCCTGAGAG 5280
QY 5281 AGGAAAGAAACATATCCACCCMWGTGACATGGTATTAAGCAAGTCCCTTCCTATAT 5340
DB 5281 AGGAAAGAAACATATCCACCCMWGTGACATGGTATTAAGCAAGTCCCTTCCTATAT 5340
QY 5341 CCCCATCCCTAATACATCCTGGGAGAGACCTACCCAGTCATTTATYTAACCCCACTG 5400
DB 5341 CCCCATCCCTAATACATCCTGGGAGAGACCTACCCAGTCATTTATYTAACCCCACTG 5400
QY 5401 CGGTTAAAGTGGCTGAGATGAGTCTTGATATACACCTTGAGTCAATCTCGATAC 5460
DB 5401 CGGTTAAAGTGGCTGAGATGAGTCTTGATATACACCTTGAGTCAATCTCGATAC 5460
QY 5461 TGCAGAAAGAACCTGAAATCCAGAGACAAAGCTAGTATTCCTGTGAACCTCTAGAG 5520
DB 5461 TGCAGAAAGAACCTGAAATCCAGAGACAAAGCTAGTATTCCTGTGAACCTCTAGAG 5520
QY 5521 ATTTGCGCTGCTCTTCAACAACAACAAGAGAAAGTAATCAATAATCCCC 5580
DB 5521 ATTTGCGCTGCTCTTCAACAACAACAAGAGAAAGTAATCAATAATCCCC 5580
QY 5581 ATGGSCTCCCTTATATATTTTCTCTATATGTTSTTTTACCTSTTTCACTCTCACT 5640
DB 5581 ATGGSCTCCCTTATATATTTTCTCTATATGTTSTTTTACCTSTTTCACTCTCACT 5640
QY 5641 GCAACCCCTCCATGCGCTGTATGACCAAGTACCTCCCTYACCMAGATTTCTATGAGA 5700
DB 5641 GCAACCCCTCCATGCGCTGTATGACCAAGTACCTCCCTYACCMAGATTTCTATGAGA 5700
QY 5701 ATGACGCGTCCGGAATATTAATGACCCCATGATAGAGTCTTTSTAAGGAAACCCC 5760
DB 5701 ATGACGCGTCCGGAATATTAATGATGCCCCATGATAGAGTCTTTSTAAGGAAACCCC 5760
QY 5761 ACCTTACGTGCCCAACCCATATGCCCGGCACTGATACCTGTCGCACTCTTGATG 5820
DB 5761 ACCTTACGTGCCCAACCCATATGCCCGGCACTGATACCTGTCGCACTCTTGATG 5820
QY 5821 CATGCAATATCATTAATTTGACAGGAAATATGATTAATCTAGTGTCTGTGAGGACTT 5880
DB 5821 CATGCAATATCATTAATTTGACAGGAAATATGATTAATCTAGTGTCTGTGAGGACTT 5880
QY 5881 GGAATCACTGTGTGTGACTTACTTCAACCAAGTATGCTGATGGGGTGAAGTT 5940
DB 5881 GGAATCACTGTGTGTGACTTACTTCAACCAAGTATGCTGATGGGGTGAAGTT 5940
QY 5941 CAAGATCAGGCAAGGAAACATGTAAGAAATATCTCCCACTCACCSGGGATCAT 6000
DB 5941 CAAGATCAGGCAAGGAAACATGTAAGAAATATCTCCCACTCACCSGGGATCAT 6000
QY 6001 GGCACCTCAGGCCCTCAACAAGGATCTCTCAAACTCATGAAACCCCTCGATCC 6060
DB 6001 GGCACCTCAGGCCCTCAACAAGGATCTCTCAAACTCATGAAACCCCTCGATCC 6060
QY 6061 CATACTCGCTGATAGCCTATTTAATACACCTTCACTGGGCTCCATGAGGTCTGCGCC 6120
DB 6061 CATACTCGCTGATAGCCTATTTAATACACCTTCACTGGGCTCATGAGGTCTGCGCC 6120
QY 6121 CAATAACCTTACTTACTGTGATATGCTCTCCCTGCACTTCARCCATATGTTCAATC 6180
DB 6121 CAATAACCTTACTTACTGTGATATGCTCTCCCTGCACTTCARCCATATGTTCAATC 6180
QY 6181 CCTGTACCTGAACAAATGGAACAATTCAGACAGAAATTAACCACTTCCGTTTATGA 6240
DB 6181 CCTGTACCTGAACAAATGGAACAATTCAGACAGAAATTAACCACTTCCGTTTATGA 6240
QY 6241 GGAACCTCTTGTTCATATSTGAAATTAACCACTACCTCAACCTCACTGTGTAAATTT 6300

Db 6241 GGAAGCTTTGTTTCCAAATGGAATTAACCACTCAACCTCACTGTTGAAATTT 6300
 Qy 6301 AGCAATACATACATACACAACCACTCCCAATGATGAGTGGTAACTCTCCACACA 6360
 Db 6301 AGCAATACATACATACACAACCACTCCCAATGATGAGTGGTAACTCTCCACACA 6360
 Qy 6361 ATAGTGTGCTTACCTTCAAGAAATATTTTGTCTGTGTGATCTTCAAGCTTATCTTTG 6420
 Db 6361 ATAGTGTGCTTACCTTCAAGAAATATTTTGTCTGTGTGATCTTCAAGCTTATCTTTG 6420
 Qy 6421 AATGCTCTTCAAGATCTATGATGCTTCTCAATCTTATGCTGATGCTTATGCTGATG 6480
 Db 6421 AATGCTCTTCAAGATCTATGATGCTTCTCAATCTTATGCTGATGCTTATGCTGATG 6480
 Qy 6481 ACTGAACAAGATTTATACATGATGATGATGATGATGATGATGATGATGATGATG 6540
 Db 6481 ACTGAACAAGATTTATACATGATGATGATGATGATGATGATGATGATGATGATG 6540
 Qy 6541 CTTCCTTTTGTATAGGAGCAGAGTGTGATGATGATGATGATGATGATGATGATG 6600
 Db 6541 CTTCCTTTTGTATAGGAGCAGAGTGTGATGATGATGATGATGATGATGATGATG 6600
 Qy 6601 ACAACCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6660
 Db 6601 ACAACCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6660
 Qy 6661 GTGCGCAGCTCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 6720
 Db 6661 GTGCGCAGCTCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 6720
 Qy 6721 CCAAATGCAAGAGCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATG 6780
 Db 6721 CCAAATGCAAGAGCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATG 6780
 Qy 6781 GGGGAAGAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6840
 Db 6781 GGGGAAGAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6840
 Qy 6841 ATTCAGATGCAATACATGATGATGATGATGATGATGATGATGATGATGATGATG 6900
 Db 6841 ATTCAGATGCAATACATGATGATGATGATGATGATGATGATGATGATGATGATG 6900
 Qy 6901 CTGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6960
 Db 6901 CTGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6960
 Qy 6961 CTGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7020
 Db 6961 CTGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7020
 Qy 7021 GAAGCTGTAACTACAAATGAGAGCCCAAGATGATGATGATGATGATGATGATGATG 7080
 Db 7021 GAAGCTGTAACTACAAATGAGAGCCCAAGATGATGATGATGATGATGATGATGATG 7080
 Qy 7081 CCCCTGAGCCGCTGTAGCCCAAGATGATGATGATGATGATGATGATGATGATGATG 7140
 Db 7081 CCCCTGAGCCGCTGTAGCCCAAGATGATGATGATGATGATGATGATGATGATGATG 7140
 Qy 7141 GAGGAATCTCAGTGTGACAACTCTACATGATGATGATGATGATGATGATGATGATGATG 7200
 Db 7141 GAGGAATCTCAGTGTGACAACTCTACATGATGATGATGATGATGATGATGATGATGATG 7200
 Qy 7201 GGTGTGAGCCCACTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATG 7260
 Db 7201 GGTGTGAGCCCACTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATG 7260
 Qy 7261 GACAGAGCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7320
 Db 7261 GACAGAGCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7320
 Qy 7321 ACCAGATCCACTTTAAACAGAGGGCTTGAAGTGTGATGATGATGATGATGATGATG 7380
 Db 7321 ACCAGATCCACTTTAAACAGAGGGCTTGAAGTGTGATGATGATGATGATGATGATG 7380

Db 7321 ACCAGATCCACTTTAAACAGAGGGCTTGAAGTGTGATGATGATGATGATGATG 7380
 Qy 7381 CTCCTAAATGCTAAATTAAGGAAGAGAGGATTAAGAAATAGCAATCATATATTCG 7440
 Db 7381 CTCCTAAATGCTAAATTAAGGAAGAGAGGATTAAGAAATAGCAATCATATATTCG 7440
 Qy 7441 MTGAGAGCAGCAGAGAGAGAGCAATGATGATGATGATGATGATGATGATGATGATG 7500
 Db 7441 MTGAGAGCAGCAGAGAGAGAGCAATGATGATGATGATGATGATGATGATGATGATG 7500
 Qy 7501 CGGCAACCCCTTGGGTCCTCCCTTTGATGATGATGATGATGATGATGATGATGATG 7560
 Db 7501 CGGCAACCCCTTGGGTCCTCCCTTTGATGATGATGATGATGATGATGATGATGATG 7560
 Qy 7561 TCTATTAATCTTGACATGCR 7582
 Db 7561 TCTATTAATCTTGACATGCR 7582
 RESULT 5
 AY101582 10222 bp DNA linear PRI 11-FEB-2004
 LOCUS Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus, allele A, complete sequence.
 DEFINITION
 ACCESSION AY101582
 VERSION AY101582.1 GI:37544399
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Mallet, P., Boulton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBMED 14757826
 REFERENCE
 AUTHORS Mallet, P., Boulton, O. and Oriol, G.
 TITLE Direct Subinjection
 JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-Biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69634 cedex 07, France
 FEATURES
 source
 1..10222
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="21"
 /isolation_source="PBMC of individual 21"
 /db_xref="Taxon:9606"
 /db_xref="map:7421-q22"
 /chromosome="7"
 /map="7421-q22"
 /sex="female"
 /cell_type="PBMC"
 /note="isolated by PCR; ERVWE1 locus, allele A endogenous_virus: HERV-W"
 1..780
 join(248..811,2887..9813)
 /note="putative mRNA transcript 1; gag-pol pseudogene"
 join(248..811,7585..9813)
 /product="envelope glycoprotein"
 /note="putative mRNA transcript 2"
 join(248..811,9248..9813)
 /note="putative mRNA transcript 3; no coding region"
 7812..9428
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AA068161.1"
 /db_xref="GI:37544400"
 /translation="MALPPIFLFTVLBSFTLTAPPCCMTSSSPYQELRMORP
 GNIDAPSYSLSKGPTPTAHTMPNRCVSAVLCHNANTHTWTGKINBSCPGGLG

Oy	2585	CCAAAGCCTTAGTAAAAACCATGATACAACTCCCGGACGTAATTCCGATGAGATTGAGAGG	2644
Db	4848	CCAAAGCCTTAGTAAAAACCATGATACAACTCCCGGACGTAATTCCGATGAGATTGAGAGG	4900
Oy	2645	CACAAAAACCCAGTGAACAGTGAAGGGATTAGTGCAGAGATCTCAGGATTTATCAATGAAGGC	2704
Db	4901	TACAGAAACCCAAACAGACAGTGA -GGTTAGGCAAGATCTCAGGATTTATCAAT -GAGGC	4958
Oy	2705	CGTTGTCCTTTATATCCCAAGCTGTAACCTAGCCCTTATACGTGMAITTTCCCAATACAGA	2764
Db	4959	TGTTGTCCTCTATATAGCCAGCTGTACTAGCCCTTATACGTGTTTCCCAATACAGA	5018
Oy	2765	GGAAACAGAGTGGTTTACASTCTGAGACCTTMAAGATGCTTTCTGATCCCTGTACA	2824
Db	5019	GGAAACAGAGTGGTTTACAGTCTTGACCTTTCAGATGCTTTCTTGATCCCTGTACA	5078
Oy	2825	TCCTGACTCTCAATTCCTGTTTGCCTTTGAAGATACTTCAAAACCCATCTCAACTCAC	2884
Db	5079	TCCTGACTCTCAATTCCTGTTTGCCTTTGAAGATACTTCAAAACCCATCTCAACTCAC	5138
Oy	2885	CTGCACTRTTTTAAACCCCAAGGTTCAAGGATAGTCCCATCTATTGGCCACAGCATTAAC	2944
Db	5139	CTGCACTRTTTTAAACCCCAAGGTTCAAGGATAGTCCCATCTATTGGCCACAGCATTAAC	5198
Oy	2945	CCAAACATGAGYCARITMTACTACTGTGACACTTTCCTCCTGCTKAGTGAATTT	3004
Db	5199	CCAAACATGAGYCARITMTACTACTGTGACACTTTCCTCCTGCTKAGTGAATTT	5256
Oy	3005	ACTTTTGGYGCYRTTCAGAAACCTGTGCCATCAAGCCAACCAAGRCCTCTMAATTT	3064
Db	5257	ACTTTTGGYGCYRTTCAGAAACCTGTGCCATCAAGCCAACCAAGRCCTCTMAATTT	5316
Oy	3065	CCTGCGYACTGTGCTACAWGTTTCCAAASAPRGCTCARCTCTGCTCAACAGAGT	3124
Db	5317	CCTGCGYACTGTGCTACAWGTTTCCAAACCAAGGCTCAACCTCTCTCAACAGAGT	5376
Oy	3125	TAAATACTTAGGRCCTAATTAATTCAAAGGACCAAGGCGCTCACTGAGGAAYRTACA	3184
Db	5377	---TACTTAGGRCCTAATTAATTCAAAGGACCAAGGCGCTCACTGAGGAAYRTACA	5432
Oy	3185	GCCATTAAGGCTTATCCATCACTCYCAAAAACCTTAAGCAACTAAGRGRTTCTTGCGRT	3244
Db	5433	GCCATTAAGGCTTATCCATCACTCAATCCAAAACCTTAAGCAACTAAGRGRTTCTTGCGRT	5492
Oy	3245	AAAYAGYTTCTGCCGAATGGAATTCCTCAGGTWTGGCAATATACCAAGYCATATAMTA	3304
Db	5493	AATAGGTTTCTCCGAAAATGGAATTCCTCAGGTWTGGCAATATACCAAGYCATATAMTA	5551
Oy	3305	CASATAATTAGGAACTCGAAAGCCAAAPACCATTATATAGATGAGVAMCTGAAGYMR	3364
Db	5552	CASATAATTAGGAACTCGAAAGCCAAAPACCATTATATAGATGAGCACTGAAGTAG	5611
Oy	3365	AAGTGCTTTCAGAGCCCTTAAGAAGGCTTAAACCCAAAGYCCCAAGTGTAAAGTTGCC	3424
Db	5612	AAGTGCTTTCAGAGCCCTTAAGAAGGCTTAAACCCAAAGYCCCAAGTGTAAAGTTGCC	5657
Oy	3425	AAACRGGGCAAGACTTTSTTATATATCTACAGAAAAAAACAGAAAYAGTCTRGGAGTCC	3484
Db	5658	AAACAGGGCAAGACTTTTCTTCAATATGTCAAG -AAAAACAGAAATAGCTCTAGAGAGTCC	5716
Oy	3485	TTAAACAGATCCBAGGGAAGAGGTTGCAACVYRTGGCRVACTGASTAAGAAAAATGATG	3544
Db	5717	TTAAACAGATCCBAGGGAAGAGGTTGCAACCTTGAGCACTGACTAAGAAAAATGATG	5776
Oy	3545	TAGTGCAAAAGGTTTGRCTATCTTTTAYGGGTAGTGTGACAGTAGCAAGTXYKTAGAT	3604
Db	5777	TAGTGCAAAAGGTTTGAACCTCATTTGTTTAGGGTATGTTGGGAGTAGCAAGTXYKTAGAT	5836
Oy	3605	CTGAAGCAGTTAAATATATACAGGGBAGAGATTTACTGTGTGACATCTCATGAKGTGA	3664
Db	5837	CTGAAGCAGTTAAATATATACAGGGBAGAGATTTACTGTGTGACATCTCATGATGTGA	5896

Db	6976	AAACCTCAGGGATTTCAGTATCTAAGTCTGGATGATCTTACAGGATTGGCCAG	7035
Qy	4804	GCCTTCCCTGTAGGACAGAAAAGGCCAAGAGTAAATAAGGACATAGTTCATTAATA	4863
Db	7036	GCCTTCCCTGTAGGACAGAAAAGGCCAAGAGTAAATAAGGACATAGTTCATTAATA	7095
Qy	4864	ATTTCCAGATTGGGACTTCCCCGAGCTTACAGAGTACATATGCCCTTCCAGGCC	4923
Db	7096	ATTTCCAGATTGGGACTTCCCCGAGCTTACAGAGTACATATGCCCTTCCAGGCC	7155
Qy	4924	ACAGTAAACCAGGGATACCCAGGCGTAAAGGTAAATACATTAACATCGGCGCTGA	4983
Db	7156	ACAGTAAACCAGGGATACCCAGGCGTAAAGGTAAATACATTAACATCGGCGCTGA	7215
Qy	4984	AGGCCACAGTCTCTCAGGGAAAGGTGAGAAAATGAATGAAYACTCAAGACATCTAAAA	5043
Db	7216	AGGCCACAGTCTCTCAGGGAAAGGTGAGAAAATGAATGAAYACTCAAGACATCTAAAA	7275
Qy	5044	AAGCAAAACCAGGAAACCCACTCAATAGCCTGTTCTGTTGCTATAGCTTTAAAAA	5103
Db	7276	AAGCAAAACCAGGAAACCCACTCAATAGCCTGTTCTGTTGCTATAGCTTTAAAAA	7335
Qy	5104	ATCTGCACCTTCCCCAAAABACAGACTAGCCCATAGGAAATGCTTAAGAAAGGCC	5163
Db	7336	ATCTGCACCTTCCCCAAAABACAGACTAGCCCATAGGAAATGCTTAAGAAAGGCC	7395
Qy	5164	TTCATATCCAAATGACCTTGCTTGCTTACCAAGACAGCCACTTATGTTGACATACCT	5223
Db	7396	TTCATATCCAAATGACCTTGCTTGCTTACCAAGACAGCCACTTATGTTGACATACCT	7455
Qy	5224	CCTTAGCCAAATATCAACAAGTTCTTAAAACTTACAAGAACCTATCCTTGAGAAGG	5283
Db	7456	CCTTAGCCAAATATCAACAAGTTCTTAAAACTTACAAGAACCTATCCTTGAGAAGG	7515
Qy	5284	GAAAAGACTATTCACCCOMGTGACATGATATTGATGAAGCCCTTGCTTAATCCC	5343
Db	7516	GAAAAGACTATTCACCCOMGTGACATGATATTGATGAAGCCCTTGCTTAATCCC	7575
Qy	5344	CATCCCTAGATACATCCTGGGAAGACCTACCCAGTATTTATATACCCCACTGGG	5403
Db	7576	CATCCCTAGATACATCCTGGGAAGACCTACCCAGTATTTATATACCCCACTGGG	7635
Qy	5404	TTAAAGTGGCTGGAGTGAAGTCTTGATACATCACTTGAGTCAATCCTGATATGCG	5463
Db	7636	TTAAAGTGGCTGGAGTGAAGTCTTGATACATCACTTGAGTCAATCCTGATATGCG	7695
Qy	5464	CAAGGAACCTGAAAAATCCAGAGCAACGCTAGCTATTCCTGTGAACCTTAAGGATT	5523
Db	7696	CAAGGAACCTGAAAAATCCAGAGCAACGCTAGCTATTCCTGTGAACCTTAAGGATT	7755
Qy	5524	TGCGCTGCTCTTCAACAACAACAGAGGAAATTAATCTAAATATCCCATG	5583
Db	7756	TGCGCTGCTCTTCAACAACAACAGAGGAAATTAATCTAAATATCCCATG	7814
Qy	5584	GSCTCCCTTATCATATTTTCTCTKTAAGTSTTTTAACTSTTTTCACTCACTGCA	5643
Db	7815	GSCTCCCTTATCATATTTTCTCTKTAAGTSTTTTAACTSTTTTCACTCACTGCA	7874
Qy	5644	CCCCCTCCATGCGCGCTGTATGACAGTAGCTCCCTTACMAAGTTCTATGAGAATG	5703
Db	7875	CCCCCTCCATGCGCGCTGTATGACAGTAGCTCCCTTACMAAGTTCTATGAGAATG	7934
Qy	5704	CAGGCTCCGGAATAATGATGACCCCATGATATAGAGCTTTTAAAGGAACCCCAAC	5763
Db	7935	CAGGCTCCGGAATAATGATGACCCCATGATATAGAGCTTTTAAAGGAACCCCAAC	7994
Qy	5764	TTCACTGCCACACCCATATGCGCGCAACTGCTATCACTTGCCACTCTTTGATGAT	5823
Db	7995	TTCACTGCCACACCCATATGCGCGCAACTGCTATCACTTGCCACTCTTTGATGAT	8054
Qy	5824	GCAATTACTCATTAATTGGACAGGAAAAATGATTAATCCTAGTTGCTTGAGACTTGA	5883

Db	8055	GCAAACTACTATTATTGGACAGAGAAAAATGATTAACTCTAGTTGTCCTGGAGACCTTGGAA	8114
QY	5884	GTCACTGTCTGTGGACTTAATTCAACCCAACTGGAATGTCTGAATGGGGGTGAAGTTCAA	5943
Db	8115	GTCACTGTCTGTGACTTACTTCAACCCAACTGGAATGTCTGAATGGGGGTGAAGTTCAA	8174
QY	5944	GATCAGGCAAGAGAAAAACATGTATAAAGAAAGTAATCTCCCACTCACCSSGGTTCATGGC	6003
Db	8175	GATCAGGCAAGAGAAAAACATGTATAAAGAAAGTAATCTCCCACTCACCOCGGGTTCATGGC	8234
QY	6004	ACCTCTAGCCCCCTTCAAAAGCACTGATCTCTCAAAACTACATGAAACCCCTCCGTACCCAT	6063
Db	8235	ACCTCTAGCCCCCTTCAAAAGCACTGATCTCTCAAAACTACATGAAACCCCTCCGTACCCAT	8294
QY	6064	ACTGCGCTGTGAAGCCTATTATTAATACCAACCTCACTGAGGCTCCATGAGAGTCTCGGCCAA	6123
Db	8295	ACTGCGCTGTGAAGCCTATTATTAATACCAACCTCACTGAGGCTCCATGAGAGTCTCGGCCAA	8354
QY	6124	AACCTTACTAATCTGTGGATATGCTCTCCCTCGAATTCARGCCATATGTTTCAATCCCT	6183
Db	8355	AACCTTACTAATCTGTGGATATGCTCTCCCTCGAATTCARGCCATATGTTTCAATCCCT	8414
QY	6184	GTACCTGAACATGGAACAATTGAGACAGAAATAAACACACTTCGGTTTTAGTAGGA	6243
Db	8415	GTACCTGAACATGGAACAATTGAGACAGAAATAAACACACTTCGGTTTTAGTAGGA	8474
QY	6244	CCTCTGTCTTCCAAATGTGAAATAAACCCATACCTCAAACTCTGTGTATAATTTAGC	6303
Db	8475	CCTCTGTCTTCCAAATGTGAAATAAACCCATACCTCAAACTCTGTGTATAATTTAGC	8534
QY	6304	AATCTACATATACACAACAACCTCCCAATGATCAGATGGGTTAACTCTCCCAACAATA	6363
Db	8535	AATCTACATATACACAACAACAACCTCCCAATGATCAGATGGGTTAACTCTCCCAACAATA	8594
QY	6364	GTCTGCGCCATCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCGCTATGCTGTTTGAAT	6423
Db	8595	GTCTGCGCCATCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCGCTATGCTGTTTGAAT	8654
QY	6424	GAGCTCTTCAGAAATCATATGTCCTTCTCTCAATTTCTAGTGGCCCCATGRCATCTACACT	6483
Db	8655	GAGCTCTTCAGAAATCATATGTCCTTCTCTCAATTTCTAGTGGCCCCATGRCATCTACACT	8714
QY	6484	GAAACAATTTATACATATATGTTCATCTTAAGCCCCCGGCAACAAAAGATACCCATCTT	6543
Db	8715	GAAACAATTTATACATATATGTTCATCTTAAGCCCCCGGCAACAAAAGATACCCATCTT	8774
QY	6544	CCTTTGTATATAGAGAGAGAGATGCTGATGGCACTAGAGTCTGGCATTTGGCGGATACACA	6603
Db	8775	CCTTTGTATATAGAGAGAGAGATGCTGATGGCACTAGAGTCTGGCATTTGGCGGATACACA	8834
QY	6604	ACCTCTACTAGTTCTCTACTCAAACTATCTCAAGAACTAATGGGGACATGGAACGGGTC	6663
Db	8835	ACCTCTACTAGTTCTCTACTCAAACTATCTCAAGAACTAATGGGGACATGGAACGGGTC	8894
QY	6664	GCCGACTTCCTGGTCACTTGGCAAGATCAACTTAATCTCTAGCAGCATATGTCCTTCRA	6723
Db	8895	GCCGACTTCCTGGTCACTTGGCAAGATCAACTTAATCTCTAGCAGCATATGTCCTTCRA	8954
QY	6724	AATGGAAGACTTTAGACTGCTGATTAACCGCATGABAGAGGGGGAAACCTGTTATTTTAAAGG	6783
Db	8955	AATGGAAGACTTTAGACTGCTGATTAACCGCATGABAGAGGGGGAAACCTGTTATTTTAAAGG	9014
QY	6784	GAAAGATGCTGTTATTTATGTTAATCAATCCGGAATGCTCACTGAGAAAGTTTAAAGAAATT	6843
Db	9015	GAAAGATGCTGTTATTTATGTTAATCAATCCGGAATGCTCACTGAGAAAGTTTAAAGAAATT	9074
QY	6844	CSAGATGGAATCAACGTATAAGCAGABAGAGCTTCAAAACACTGGAACCTTGGGGCTCTCTC	6903
Db	9075	CSAGATGGAATCAACGTATAAGCAGABAGAGAGCTTCAAAACACTGGAACCTTGGGGCTCTCTC	9134
QY	6904	AGCCRAATGGAATGCCCGGAGATTCCTCCCTCTTAAAGCACTTACAGAGCTTAATATATGCTA	6963
Db	9135	AGCCRAATGGAATGCCCGGAGATTCCTCCCTCTTAAAGCACTTACAGAGCTTAATATATGCTA	9194

QY	6954	CTCCCTTTTGAA	CCTGTATCTTTFA	CCTCCCTTGTAA	CTTGTCTCTTCCAGAA	TGAA	7023
Db	9195	CTCCCTTTTGAA <td>CCTGTATCTTTAA <td>CCTCCTTTGTAA <td>CTTGTCTCTTCCAGAA <td>TGAA</td> <td>9254</td> </td></td></td>	CCTGTATCTTTAA <td>CCTCCTTTGTAA <td>CTTGTCTCTTCCAGAA <td>TGAA</td> <td>9254</td> </td></td>	CCTCCTTTGTAA <td>CTTGTCTCTTCCAGAA <td>TGAA</td> <td>9254</td> </td>	CTTGTCTCTTCCAGAA <td>TGAA</td> <td>9254</td>	TGAA	9254
QY	7024	GCTGTAAATCTA <td>CAATGAGAGCC</td> <td>CAAGATGCA <td>AGACTTAACG</td> <td>CGAGACC</td> <td>7083</td> </td>	CAATGAGAGCC	CAAGATGCA <td>AGACTTAACG</td> <td>CGAGACC</td> <td>7083</td>	AGACTTAACG	CGAGACC	7083
Db	9255	GCTGTAAATCTA <td>CAATGAGAGCC</td> <td>CAAGATGCA <td>AGACTTAACG</td> <td>CGAGACC</td> <td>9314</td> </td>	CAATGAGAGCC	CAAGATGCA <td>AGACTTAACG</td> <td>CGAGACC</td> <td>9314</td>	AGACTTAACG	CGAGACC	9314
QY	7084	CTGACCGGGCT <td>GTAGGCCA <td>CGATCTGATGTTAA <td>TGACATCAAGGAC</td> <td>CCCTCTGAG</td> <td>7143</td> </td></td>	GTAGGCCA <td>CGATCTGATGTTAA <td>TGACATCAAGGAC</td> <td>CCCTCTGAG</td> <td>7143</td> </td>	CGATCTGATGTTAA <td>TGACATCAAGGAC</td> <td>CCCTCTGAG</td> <td>7143</td>	TGACATCAAGGAC	CCCTCTGAG	7143
Db	9315	CTGACCGGGCT <td>GTAGGCCA <td>CGATCTGATGTTAA <td>TGACATCAAGGAC</td> <td>CCCTCTGAG</td> <td>9374</td> </td></td>	GTAGGCCA <td>CGATCTGATGTTAA <td>TGACATCAAGGAC</td> <td>CCCTCTGAG</td> <td>9374</td> </td>	CGATCTGATGTTAA <td>TGACATCAAGGAC</td> <td>CCCTCTGAG</td> <td>9374</td>	TGACATCAAGGAC	CCCTCTGAG	9374
QY	7144	GAAATCTCAG <td>CTGCACAACT <td>CTTAAGCCCA <td>TTTGAAGAGAGG</td> <td>AGTGAAGCGGT</td> <td>7203</td> </td></td>	CTGCACAACT <td>CTTAAGCCCA <td>TTTGAAGAGAGG</td> <td>AGTGAAGCGGT</td> <td>7203</td> </td>	CTTAAGCCCA <td>TTTGAAGAGAGG</td> <td>AGTGAAGCGGT</td> <td>7203</td>	TTTGAAGAGAGG	AGTGAAGCGGT	7203
Db	9375	GAAATCTCAG <td>CTGCACAACT <td>CTTAAGCCCA <td>TTTGAAGAGAGG</td> <td>AGTGAAGCGGT</td> <td>9434</td> </td></td>	CTGCACAACT <td>CTTAAGCCCA <td>TTTGAAGAGAGG</td> <td>AGTGAAGCGGT</td> <td>9434</td> </td>	CTTAAGCCCA <td>TTTGAAGAGAGG</td> <td>AGTGAAGCGGT</td> <td>9434</td>	TTTGAAGAGAGG	AGTGAAGCGGT	9434
QY	7204	SGTGGGCAAC <td>CTCCCAAC</td> <td>AGACCTTGA <td>GGTTTCTGTGAGAT <td>GGGGAGCTGAGAGAC</td> <td>7263</td> </td></td>	CTCCCAAC	AGACCTTGA <td>GGTTTCTGTGAGAT <td>GGGGAGCTGAGAGAC</td> <td>7263</td> </td>	GGTTTCTGTGAGAT <td>GGGGAGCTGAGAGAC</td> <td>7263</td>	GGGGAGCTGAGAGAC	7263
Db	9435	SGTGGGCAAC <td>CTCCCAAC</td> <td>AGACCTTGA <td>GGTTTCTGTGAGAT <td>GGGGAGCTGAGAGAC</td> <td>9494</td> </td></td>	CTCCCAAC	AGACCTTGA <td>GGTTTCTGTGAGAT <td>GGGGAGCTGAGAGAC</td> <td>9494</td> </td>	GGTTTCTGTGAGAT <td>GGGGAGCTGAGAGAC</td> <td>9494</td>	GGGGAGCTGAGAGAC	9494
QY	7264	AGACTTAAGCT <td>GTATTTCTT <td>AGAGCTGATTA <td>AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>7323</td> </td></td></td>	GTATTTCTT <td>AGAGCTGATTA <td>AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>7323</td> </td></td>	AGAGCTGATTA <td>AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>7323</td> </td>	AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>7323</td>	AGTGGAGAGTACC	7323
Db	9495	AGACTTAAGCT <td>GTATTTCTT <td>AGAGCTGATTA <td>AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>9554</td> </td></td></td>	GTATTTCTT <td>AGAGCTGATTA <td>AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>9554</td> </td></td>	AGAGCTGATTA <td>AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>9554</td> </td>	AGATCTTAAGCTTA <td>AGTGGAGAGTACC</td> <td>9554</td>	AGTGGAGAGTACC	9554
QY	7324	ACATCCACCTT <td>TAAACAGGGG</td> <td>CTTGCACTTA <td>GTCAACCTGACCA</td> <td>TACAGAGGCTC</td> <td>7383</td> </td>	TAAACAGGGG	CTTGCACTTA <td>GTCAACCTGACCA</td> <td>TACAGAGGCTC</td> <td>7383</td>	GTCAACCTGACCA	TACAGAGGCTC	7383
Db	9555	ACATCCACCTT <td>TAAACAGGGG</td> <td>CTTGCACTTA <td>GTCAACCTGACCA</td> <td>TACAGAGGCTC</td> <td>9614</td> </td>	TAAACAGGGG	CTTGCACTTA <td>GTCAACCTGACCA</td> <td>TACAGAGGCTC</td> <td>9614</td>	GTCAACCTGACCA	TACAGAGGCTC	9614
QY	7384	ACTAAATAGCT <td>TAATTAAGCA</td> <td>AGAGAGGTA</td> <td>GAATAGCAAT</td> <td>CAATATATTTGCMGT</td> <td>7443</td>	TAATTAAGCA	AGAGAGGTA	GAATAGCAAT	CAATATATTTGCMGT	7443
Db	9615	ACTAAATAGCT <td>TAATTAAGCA</td> <td>AGAGAGGTA</td> <td>GAATAGCAAT</td> <td>CAATATATTTGCMGT</td> <td>9674</td>	TAATTAAGCA	AGAGAGGTA	GAATAGCAAT	CAATATATTTGCMGT	9674
QY	7444	AGAGACAGCAG	AGAGAGGCA	CAATGATTCGGGAT	TTAAACCAAT	GTTCGAGCGGCAAGG	7503
Db	9675	AGAGACAGCAG	AGAGAGGCA	CAATGATTCGGGAT	TTAAACCAAT	GTTCGAGCGGCAAGG	9734
QY	7504	CAACCCCTTTG	GGGTCCTCC	CTTTGTATG	GAGACTGTGTTT	CATGCTATTTCACTCT	7563
Db	9735	CAACCCCTTTG	GGGTCCTCC	CTTTGTATG	GAGACTGTGTTT	CATGCTATTTCACTCT	9794
QY	7564	ATTAAATCTT	GCARCTGCR	7582			
Db	9795	ATTAAATCTT	GCARCTGCA	9813			

RESULT 6

AY101585

LOCUS

DEFINITION Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVW1 locus.

ACCESSION AY101585

VERSION AY101585.1

KEYWORDS GI:37544405

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

1 (bases 1 to 10222) Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L., and Mandrand, B. The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

2 (bases 1 to 10222) Mallet, F., Bouton, O., and Oriol, G. Direct Submission Submitted (06-MAY-2002) Retrovirology Departement, UMR 2142 CNRS-biomeieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

Location/Qualifiers

1..10222

Query Match	85.0%; Score 6448; DB 9; Length 10222;	Query Match
Best Local Similarity 95.7%; Pred. No. 0;	Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;	Query Match
QY 606 GATGGGAAAGCTTCCCGCAAGCAAAAAGCCCTTAAGACGATTCTGGARATTGGGA 665	QY 606 GATGGGAAAGCTTCCCGCAAGCAAAAAGCCCTTAAGACGATTCTGGARATTGGGA 665	QY 606 GATGGGAAAGCTTCCCGCAAGCAAAAAGCCCTTAAGACGATTCTGGARATTGGGA 665
DB 2886 GATGGGAAAGCTTCCCGCAAGCAAAAAGCCCTTAAGACGATTCTGGARATTGGGA 2945	DB 2886 GATGGGAAAGCTTCCCGCAAGCAAAAAGCCCTTAAGACGATTCTGGARATTGGGA 2945	DB 2886 GATGGGAAAGCTTCCCGCAAGCAAAAAGCCCTTAAGACGATTCTGGARATTGGGA 2945
QY 666 MCATTGGACCCCTCAGACCTAAGAAAGACGACTATATTCTTCTGCAAGTCCGCGCTG 725	QY 666 MCATTGGACCCCTCAGACCTAAGAAAGACGACTATATTCTTCTGCAAGTCCGCGCTG 725	QY 666 MCATTGGACCCCTCAGACCTAAGAAAGACGACTATATTCTTCTGCAAGTCCGCGCTG 725
DB 2946 CCAATTGGACCCCTCAGACCTAAGAAAGACGACTATATTCTTCTGCAAGTCCGCGCTG 3005	DB 2946 CCAATTGGACCCCTCAGACCTAAGAAAGACGACTATATTCTTCTGCAAGTCCGCGCTG 3005	DB 2946 CCAATTGGACCCCTCAGACCTAAGAAAGACGACTATATTCTTCTGCAAGTCCGCGCTG 3005
QY 726 GCACTCTCGAGGGAAGTAAATTAATTAACAACCATTTACAGCTAGACTCTTTTGTGAA 785	QY 726 GCACTCTCGAGGGAAGTAAATTAATTAACAACCATTTACAGCTAGACTCTTTTGTGAA 785	QY 726 GCACTCTCGAGGGAAGTAAATTAATTAACAACCATTTACAGCTAGACTCTTTTGTGAA 785
DB 3006 GCACTCTCGAGGGAAGTAAATTAATTAACAACCATTTACAGCTAGACTCTTTTGTGAA 3065	DB 3006 GCACTCTCGAGGGAAGTAAATTAATTAACAACCATTTACAGCTAGACTCTTTTGTGAA 3065	DB 3006 GCACTCTCGAGGGAAGTAAATTAATTAACAACCATTTACAGCTAGACTCTTTTGTGAA 3065
QY 786 AAGCAAAATGAGTGAAGTGCATTAAGTACAAATCTTTTCAATTAGAACAACCTAC 845	QY 786 AAGCAAAATGAGTGAAGTGCATTAAGTACAAATCTTTTCAATTAGAACAACCTAC 845	QY 786 AAGCAAAATGAGTGAAGTGCATTAAGTACAAATCTTTTCAATTAGAACAACCTAC 845
DB 3066 AAGCAAAATGAGTGAAGTGCATTAAGTACAAATCTTTTCAATTAGAACAACCTAC 3125	DB 3066 AAGCAAAATGAGTGAAGTGCATTAAGTACAAATCTTTTCAATTAGAACAACCTAC 3125	DB 3066 AAGCAAAATGAGTGAAGTGCATTAAGTACAAATCTTTTCAATTAGAACAACCTAC 3125
QY 846 AATTATGTAATAAAGTGAATTATAGCCCTCAGGAACCTTACAGAGTCTACCTCCATAC 905	QY 846 AATTATGTAATAAAGTGAATTATAGCCCTCAGGAACCTTACAGAGTCTACCTCCATAC 905	QY 846 AATTATGTAATAAAGTGAATTATAGCCCTCAGGAACCTTACAGAGTCTACCTCCATAC 905
DB 3126 AATTATGTAATAAAGTGAATTATAGCCCTCAGGAACCTTACAGAGTCTACCTCCATAC 3185	DB 3126 AATTATGTAATAAAGTGAATTATAGCCCTCAGGAACCTTACAGAGTCTACCTCCATAC 3185	DB 3126 AATTATGTAATAAAGTGAATTATAGCCCTCAGGAACCTTACAGAGTCTACCTCCATAC 3185
QY 906 CCAAGCATCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCAA 965	QY 906 CCAAGCATCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCAA 965	QY 906 CCAAGCATCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCAA 965
DB 3186 CCAAGCATCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCAA 3245	DB 3186 CCAAGCATCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCAA 3245	DB 3186 CCAAGCATCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCAA 3245
QY 966 AAGAGATTAACAACAAGGTAAACAGTGAACCAAGAAGTGCATTAATCCCAATTATGA 1025	QY 966 AAGAGATTAACAACAAGGTAAACAGTGAACCAAGAAGTGCATTAATCCCAATTATGA 1025	QY 966 AAGAGATTAACAACAAGGTAAACAGTGAACCAAGAAGTGCATTAATCCCAATTATGA 1025
DB 3246 AAGAGATTAACAACAAGGTAAACAGTGAACCAAGAAGTGCATTAATCCCAATTATGA 3305	DB 3246 AAGAGATTAACAACAAGGTAAACAGTGAACCAAGAAGTGCATTAATCCCAATTATGA 3305	DB 3246 AAGAGATTAACAACAAGGTAAACAGTGAACCAAGAAGTGCATTAATCCCAATTATGA 3305
QY 1026 CCCCTCCAGCAGTGGAGAAAGAAATTCGCGCCAGCAGAGTGCATGTCCTTTTAT 1085	QY 1026 CCCCTCCAGCAGTGGAGAAAGAAATTCGCGCCAGCAGAGTGCATGTCCTTTTAT 1085	QY 1026 CCCCTCCAGCAGTGGAGAAAGAAATTCGCGCCAGCAGAGTGCATGTCCTTTTAT 1085

```

Db      3306  CCCCT-CCAGCATGGGAGAGAGATTGCGCCCAAGCCAGATGCACTGCTTTTC 3364
Qy      1086  TCTCCAGACTTTAAAGCAATTAAGAGACTTAGTAATTTCTGATTAATCTGATG 1145
Db      3365  TCTCCAGACTTTAAAGCAATTAAGAGACTTAGTAATTTCTGATTAATCTGATG 3424
Qy      1146  CTATTTTBTGTTTAAAGAGGTTAGACAAATTTCTTGTATCTGACATGAGATATTA 1205
Db      3425  CTATTTTBTGTTTAAAGAGGTTAGACAAATTTCTTGTATCTGACATGAGATATTA-A 3483
Qy      1206  TGTCACTGCTAAATCAGACACTTAACCCCAATGAGAGAGTGCACCAATTAATGAGCT 1265
Db      3484  TGTCACTGCTAAATCAGACACTTAACCCCAATGAGAGAGTGCACCAATTAATGAGCT 3543
Qy      1266  GAGGTGTTGGCGATCTGTGATCTCAGTCAAGTCAATGATGATGATGACAAAGAG 1325
Db      3544  GAGGTGTTGGCGATCTGTGATCTCAGTCAAGTCAATGATGATGATGACAAAGAG-AGG 3600
Qy      1326  AAAGAAATGATTTCCCAAGGCGCAGACAGGCTTCCAGTCTASACCTCATTTGGGAC 1385
Db      3601  AAAGAAATGATTTCCCAAGGCGCAGACAGGCTTCCAGTCTASACCTCATTTGGG- 3657
Qy      1386  ACAGAAATCAGTAATGAGAGATGAGTGCAGACACTTGTCTAATCTTGTGCTASAA 1445
Db      3658  ACAGAAATCAGTAATGAGAGATGAGTGCAGACACTTGTCTAATCTTGTGCTASAA 3716
Qy      1446  GGAATTAAGAAATCTASAGAAATCTAYGATTAATCTCAATGATGCTCAATACACA 1505
Db      3717  GGAATTAAGAAATCTASAGAG-AGTCTATGATTAATCTCAATGATGCTCAATACACA 3775
Qy      1506  GGGAGAGGAAAGAAATCTCTACTGCTTTCTGAGAGACTTAAGAGGCAATTAAGAG 1565
Db      3776  -GGGAGAGGAAAGAAATCTCTACTGCTTTCTGAGAGACTTAAGAGGCAATTAAGAGAGC 3834
Qy      1566  GTGCTCTCTGTCACTGACTCTTCTGAAGGCAACTATCTTAAGGTAAGTTATCA 1625
Db      3835  GTGCTCTCTGTCACTGACTCTTCTGAAGGCAACTATCTTAAGGTAAGTTATCA 3894
Qy      1626  CTCAGTCACTGACAGACTTAG-AAAAAATTCAAAAGTCTGCGTAGGCCCGAGCAAA 1684
Db      3895  CTCAGTCACTGACAGACTTAGAATAAAAAAATTCAAAAGTCTGCGTAGGCCCGAGCAAA 3954
Qy      1685  ACTTAAGAAATCCCTATTGAATCTTGCACTGTGGTTTTTATTAATGAGATCAGAGAGC 1744
Db      3955  ACTTAAGAAATCCCTATTGAATCTTGCACTGTGGTTTTTATTAATGAGATCAGAGAGC 4014
Qy      1745  AGGCGAAACAGAGCAAAAGGGATTAAAAAAGGCAACCGCTTAGTATGATGACCTCAGG 1804
Db      4015  AGGCGAAACAGAGCAAAAGGGATTAAAAAAGGCAACCGCTTAGTATGATGACCTCAGG 4074
Qy      1805  CAAATGACTTTGAGGCTCTGAAAAGGAAAAGTGGGCAAAATTGATGCTTAATAG 1864
Db      4075  CAAATGACTTTGAGGCTCTGAAAAGGAAAAGTGGGCAAAATTGATGCTTAATAG 4134
Qy      1865  GCTTGCTTCAAGTGGCTTAACAAGACACTTTAAAAAAGTGTCCAAAGTAGAG 1924
Db      4135  GCTTGCTTCAAGTGGCTTAACAAGACACTTTAAAAAAGTGTCCAAAGTAGAG 4194
Qy      1925  CCGGCTTGTGTCATGCGCTTTATTCAAGGGAATCACTGAAAGGCCCATGCCCCAGG 1984
Db      4195  CCGGCTTGTGTCATGCGCTTTATTCAAGGGAATCACTGAAAGGCCCATGCCCCAGG 4254
Qy      1985  GGAAGAAAGTCTTTGAGTGAAGCACTTAACAGATGATCCAGAGCAAGAGCTGAGG 2044
Db      4255  GGAAGAAAGTCTTTGAGTGAAGCACTTAACAGATGATCCAGAGCAAGAGCTGAGG 4314
Qy      2045  TGCTTGAGGAGAGCGCATCCATGCACTACCTTCAAGAGCCCTGGGTATGCTTAAC 2104
Db      4315  TGCTTGAGGAGAGCGCATCCATGCACTACCTTCAAGAGCCCTGGGTATGCTTAAC 4374
Qy      2105  ATTGAAGGCAAGAGGTTGCTCTGAGACATGCTGGGCTTCTTAATGCTTCTTC 2164
Db      4375  ATTGAAGGCAAGG-AGGTTGCTCTGAGACATGCTGGGCTTCTTAATGCTTCTTC 4433

```

```

Qy      2165  TGTCCGAGCACTGCTCTCCAGATCTGTCACTAATTTGAGGGGCTCCNTAAGACGAGCA 2224
Db      4434  TGTCCGAGCACTGCTCTCCAGATCTGTCACTA-TGTGAGGGGCTCC-TAAGAGGAGCA 4491
Qy      2225  GTCATAGATCTTTTCCCAAGCACTAAGTTATGAACTGGGAGCTTTATTTCTTCAAC 2284
Db      4492  GTCATAGATCTTTTCCCAAGCACTAAGTTATGAACTGGGAGCTTTATTTCTTCAAC 4549
Qy      2285  ATGCTTTCTAATTAATGTTGAAGGCCCATCTACTGTTTGAAGGAGACAAATTCAGCA 2344
Db      4550  ATGCTTTCTAATTAATGTTGAAGGCCCATCTACTGTTTGAAGGAGACAAATTCAGCA 4609
Qy      2345  AAGCAGGGGCAATTAACCTGAACTAGAGAGAGCAACCGTTGTTGTTGCCCTG 2404
Db      4610  AAGCAGGGGCAATTAACCTGAACTAGAGAGAGCAACCGTTGTTGTTGCCCTG 4668
Qy      2405  CTTGAGAGAGAAATTAATCTGAAAGTGTGGCAACAGAGGCAATATGACGAGCCAAA 2464
Db      4669  CTTGAGAGAGAAATTAATCTGAAAGTGTGGCAACAGAGGCAATATGACGAG-CAAA 4727
Qy      2465  GAATGCCGCTGCTGCAAGTTAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA 2524
Db      4728  GAATGCCGCTGCTGCTGCAAGTTAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA 4787
Qy      2525  GTACCCCTGAGACCCCAAGGCCCAAGAGATTCAAAAGATTGTAAAGACTTAAAGC 2584
Db      4788  GTACCCCTGAGACCCCAAGGCCCAAGAGATTCAAAAGATTGTAAAGACTTAAAGC 4847
Qy      2585  CCAAGCTTGAATAACATGATGATTAATCTCCCTGAGTAAATTCCTGATGATGAGAG 2644
Db      4848  CCAAGCTTGAATAACATGATGATTAATCTCCCTGAGTAAATTCCTGATGATGAGAG 4900
Qy      2645  CACAGAAACCAAGGACAGTGAAGGTTGTGCAAGATCTCAGATTAATCAATGAGAGC 2704
Db      4901  TACAGAAACCAAGGACAGTGAAG-GGTATGTCAGATCTCAGATTAATCAATGAGAGC 4958
Qy      2705  CGTTGCTCTTATTAACCAAGCTGATCTTAAGCTTAACTGTGATGATTTCCAAATACAGA 2764
Db      4959  TGTGTTCTCTTATTAACCAAGCTGATCTTAAGCTTAACTGTGATGATTTCCAAATACAGA 5018
Qy      2765  GGAAGCAGATGTTTACATCTCTGACCTTTAGAGATGCTTCTTCTGCAATCCTGATCA 2824
Db      5019  GGAAGCAGATGTTTACATCTCTGACCTTTAGAGATGCTTCTTCTGCAATCCTGATCA 5078
Qy      2825  TCTGACTCTCAATCTGTTGCTTGAAGATTAATCTTCAAAACCAATCTCAATCAGC 2884
Db      5079  TCTGACTCTCAATCTGTTGCTTGAAGATTAATCTTCAAAACCAATCTCAATCAGC 5138
Qy      2885  CTGACTTCTTAAACCCCAAGGTTCAAGAGTAGTCCCATCTAATTTGGCCAGGACTTAGC 2944
Db      5139  CTGACTTCTTAAACCCCAAGGTTCAAGAGTAGTCCCATCTAATTTGGCCAGGACTTAGC 5198
Qy      2945  CCAAGACTTGAATCAATTAATCACTGACACTCTTGTCTTCTGATGATGATTT 3004
Db      5199  CCAAGACTTGAATCAATTAATCACTGACACTCTTGTCTTCTGATGATGATTT 5256
Qy      3005  ACTTTTGGCCGCTTGAAGAACTTGGGCACTGAAGCAACCAACGCTTMAATTT 3064
Db      5257  ACTTTTGGCCGCTTGAAGAACTTGGGCACTGAAGCAACCAACGCTTMAATTT 5316
Qy      3065  CCTGCACTGTTGAGTCAAGGTTTCAAAAGGAGAGAGTCACTGCTCAAGAGAGT 3124
Db      5317  CCTGCACTGTTGAGTCAAGGTTTCAAAAGGAGAGAGTCACTGCTCAAGAGAGT 5376
Qy      3125  TAAATCTTGAAGTAAATTAATCAAGGCAACAGGCTTCACTGAGAGATTAATCA 3184
Db      5377  ---TAAATCTTGAAGTAAATTAATCAAGGCAACAGGCTTCACTGAGAGATTAATCA 5432
Qy      3185  GCTTATACAGGCTTATCTCATCTCAATCAAAACCTTAAAGCACTAAGGATTTCTTGGCT 3244
Db      5433  GCTTATACAGGCTTATCTCATCTCAATCAAAACCTTAAAGCACTAAGGATTTCTTGGCT 5492

```

QY	3245	AAVAGYTTCCCGGAAATGATTC	CCGAGTWTGGRAATGACAGCATTAATA	3304
Db	5493	AATAGTTTCGCGAAATGATTT	CCGAGTATGGGAAATGACAGCATTAATA	5551
QY	3305	CASPAATTAAGGAAATCAGAAAGCAATA	CCCATTTTARTAAAGTGAAYAACTGAAGYMR	3364
Db	5552	CACHTATTAAGGAAATCAGAAAGCAATA	CCCATTTTARTAAAGTGAACAACTGAATAG	5611
QY	3365	AAGTGGCTTTCAGAGCCCTTAAGAGGC	CTTAAGCCCAAGYCCCAAGTTTAAAGTTGCC	3424
Db	5612	AAAGTGGCTTTCAGAGCCC-----	TTAACCAAGCCCAAGTTTAAAGTTGCC	5657
QY	3425	AACGGGGCAAACTTTTSTYATATRTCA	AGAAAAAACAAGAAAYAGCTCTRGATCC	3484
Db	5658	AACAGGGGCAAACTTTTCTCATATGTCA	AG-AAAAACAGAAATAGCTCTTAGAATCC	5716
QY	3485	TTACACAGTCTCRRAGGGAYGAGCTT	GCACACRYRGGGRYACCTGASPAAGGAAAYTATG	3544
Db	5717	TTACACAGATCCGAGGGAYGAGCTT	GCACCTGTGGCATACCTGATCTTAAGGAAATTAATG	5776
QY	3545	TAGTGGCAAAAGGTGTGRCTTCATT	GTTTTAYGGTATGTGTGGCAGTAGCAATYKATAT	3604
Db	5777	TAGTGGCAAAAGGTGTGACCTCATTT	GTTTTAYGGTATGTGTGGCAGTGTGATAT	5836
QY	3605	CTGAAGCATTTAAATAATATACAGGG	RAGAATCTTACTGTGTGGCATCTCATAKTGA	3664
Db	5837	CTGAAGCATTTAAATAATATACAGGG	AAAGATCTTACTGTGTGACATCTCATATATGTA	5896
QY	3665	AYRCATPCTACCTGTCTTAAGGAGC	CTGTGGCTGTGACAACACGTTTACTTAARTC	3724
Db	5897	ATGCAATPCTACCTGTCTTAAGGAGC	CTGTGGCTGTGACAACCTGTTTACTTAARTC	5956
QY	3725	AGGCTCTATTACTTGAARGGCCAGT	GTCTGCRACCTGTGCAACTTGTAAACCAG	3784
Db	5957	AGGCTCTATTACTTGAARGGCCAGT	GTCTGCGACCTGTGCACTTGTGCAACTTTAAACCAG	6016
QY	3785	YCNCAATTTCTTCAGACCAATGAG	AAAAAGATAPAAVATATCTGTCAACAAATATTTCTC	3844
Db	6017	CCACATTTCTTCAGACCAATGAG	AAAAAGATAPAAACATATCTGTCAACAAATATTTCTC	6076
QY	3845	AAACCTATGCACTTCGAGGGACCTT	YAGARGTTCCTTACATGATCCYACCTTCAAC	3904
Db	6077	AAACCTATGCACTTCGAGGGACCTT	YAGAGTTCCTTACATGATCCCAAC-TCAC	6135
QY	3905	TTGTATATCTGATGGAAGTTCTCTT	GTGTAAGAAAAAGACTTCAAAAAGYGGGATATGACAGTG	3964
Db	6136	TTGTATATCTGATGGAAGTTCTCTT	GTGTAAGAAAAAGACTTCBAAAAAGTGGGATATGACAGTG	6195
QY	3965	GTCACTGATTAATGGAATAATTGA	AAAGTATCCCTCACTCCAGACATAAGTCTYAGCTR	4024
Db	6196	GTCACTGATTAATGGAATAACTTGA	AAAGTATCCCTCACTCCAGACATAAGTCTCACTGA	6255
QY	4025	GCAGAACTTAATAGCCCTCAATYK	GGGCACTAGAAATTAGAGAAAGRAAAAGGGVAAATTA	4084
Db	6256	GCAGAACTTAATAGCCCTCAATYK	GGGCACTAGAAATTAGAGAAAGRAAAAGGGCAAAATTA	6315
QY	4085	TATACAGACTCTBATATGCTATCA	CTAGTCNTCCATGGCCCAATGMRGCAATATGSAAGAA	4144
Db	6316	TATACAGACTCTBAATATGCTTACT	TAAGTCTTCATGCCATGACGAATATGGAAGAA	6375
QY	4145	AGGGAATTCCTAATCTTGYAGRGA	ACACTATCAMAATCAGAAAGCCATTAGABATTA	4204
Db	6376	AGGGAATTCCTAATCTTGAAGGA	CACTATCAACATTCAGAAAGCCATTAGAAATTA	6435
QY	4205	TTATYGGCGWTCAGAAAACCTTA	BAAGGTGMAAGCTTAACACTGCGYGGGGTCATCANAA	4264
Db	6436	TTATYGGCGWTCAGAAAACCTTA	BAAGGTGMAAGCTTACACTGCGGGGTCTATCAGAA	6495
QY	4265	GGAAAGRAAAAGGGAATASAA	GRGAAYTGCACAGAKATATTGGAACMAAAAGAGCTGCA	4324
Db	6496	GGAAAGRAAAAGGGAATATGAG	AGAACTGCCACAGATATTGGAAGCCMAAAGAGCTGCA	6555
QY	4325	AGGCAAGACCTCCATTAAGAAAT	GTCTTAATACTCCCTTAAGTATAGGCTAATCCCTTC	4384

D	b	6556	AGCGAGAGACCTTCATTGAAATGTTTAAACCAACCCTGATATAGGGTATCCCTC	6615
Q	y	4385	CGGGAAACCAAGCCCCAGTACTCAGCAGAGAAACAGATGGGGAACTTCAGAG- CAG	4443
D	b	6616	CGGGAAACCAAGCCCCAGTACTCAGCAGAGAAACAGATGGGGAACTTCAGAGACAG	6675
Q	y	4444	TTTTCTCCCTCCGGGACGGTTAGGACCTGGAAGAAGGAAAAATATCTTTGCGTGCAT	4503
D	b	6676	TTTTCTCCCTCCGGGACGGTTAGGACCTGGAAGAAGGAAAAATATCTTTGCGTGCAT	6735
Q	y	4504	CCAAATGGAATTTACTTAAACCCCTTCATCAAACTTTCACTTAGCATGATAGACCCA	4563
D	b	6736	CCAAATGGAATTTACTTAAACCCCTTCATCAAACTTTCACTTAGCATGATAGACCCA	6795
Q	y	4564	TCAGATGGCCAAATCATTTATTTACTGACACAGGCTTTTCAAACTATCAAGCATATKT	4623
D	b	6796	TCAGATGGCCAAATCATTTATTTACTGACACAGGCTTTTCAAACTATCAAGCATATKT	6855
Q	y	4624	CAGGGCCCTGTGAATGTGCGCARAAATATATCCCTGCTTATTCGCAAGCTCCTCAG	4683
D	b	6856	CAGGGCCCTGTGAATGTGCGCARAAATATATCCCTGCTTATTCGCAAGCTCCTCAG	6915
Q	y	4684	ABAAACAABAAACAGGSCCATTAACCCGABAAABACTGSCAACTGATTTTACCCACAA	4743
D	b	6916	ABAAACAABAAACAGGSCCATTAACCCGABAAABACTGSCAACTGATTTTACCCACAA	6975
Q	y	4744	AAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTATATATTTACGGGTTGGCARAG	4803
D	b	6976	AAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTATATATTTACGGGTTGGCARAG	7035
Q	y	4804	GCTTTCCCTGTAGACACAGAAAGGCCAAGAGGTAAATTAAGSCATAGTTCATGAATA	4863
D	b	7036	GCTTTCCCTGTAGACACAGAAAGGCCAAGAGGTAAATTAAGSCATAGTTCATGAATA	7095
Q	y	4864	ATTCCCAATTCGACCTTCCCGAGGCTTACAGATGACATAGCCCTGTTTCAAGCC	4923
D	b	7096	ATTCCCAATTCGACCTTCCCGAGGCTTACAGATGACATAGCCCTGTTTCAAGCC	7155
Q	y	4924	ACAGTAAACCCAGGAGGTATCCAGGCGTTAGGTATCGATATCATCTACCTGCGCTGA	4983
D	b	7156	ACAGTAAACCCAGGAGGTATCCAGGCGTTAGGTATCGATATCATCTACCTGCGCTGA	7215
Q	y	4984	AGGCAACAGTCTCAGGGAGGTGAGAAATGAATGAATATCTCAAAGACATCTTAAA	5043
D	b	7216	AGGCAACAGTCTCAGGGAGGTGAGAAATGAATGAATATCTCAAAGACATCTTAAA	7275
Q	y	5044	AAGCAAAACCCAGGAAACCACTCATAGGCGTGTCTGTGCTATAGCTTAAAAAGA	5103
D	b	7276	AAGCAAAACCCAGGAAACCACTCATAGGCGTGTCTGTGCTATAGCTTAAAAAGA	7335
Q	y	5104	ATCTGCACTTTCCCAAAAACAGAGACTTAGCCCATACGAAATGCTGTATGGAAGCCC	5163
D	b	7336	ATCTGCACTTTCCCAAAAACAGAGACTTAGCCCATACGAAATGCTGTATGGAAGCCC	7395
Q	y	5164	TTCAATACCAATGACCTTGTGCTTGAACCAAGACACCACTTAGTTGACAGACATCACT	5223
D	b	7396	TTCAATACCAATGACCTTGTGCTTGAACCAAGACACCACTTAGTTGACAGACATCACT	7455
Q	y	5224	CCTTAGCCAAATATCAACAGTTCTTAAACATTAACAAGAACCTTACCTGAGAAAGG	5283
D	b	7456	CCTTAGCCAAATATCAACAGTTCTTAAACATTAACAAGAACCTTACCTGAGAAAGG	7515
Q	y	5284	GAAAAAGACTATTCACACCCMGGTGACATGTAATTAGTCAAGTCCCTTCCTTAATCCC	5343
D	b	7516	GAAAAAGACTATTCACACCCMGGTGACATGTAATTAGTCAAGTCCCTTCCTTAATCCC	7575
Q	y	5344	CATCCCTAGATCATCTCTGGGAAGACCTTACCACTATTTTATATACCCCAACGCGG	5403
D	b	7576	CATCCCTAGATCATCTCTGGGAAGACCTTACCACTATTTTATATACCCCAACGCGG	7635
Q	y	5404	TTAAAGTGCGTAGAGTGGAGTCTTGATATCATCACTTGAGTCAAACTCTGATCTGC	5463


```
Db 7636 TTAAGTGGCTGAGTGGAGCTTGGATACATCACACTTGAATCAATCTGGATACTGC 7695
Qy 5464 CAAAGAAACCTGAAAATCCAGAGACAAAGCTAGCTATTTCTGTGAACCTCTAGAGATT 5523
Db 7696 CAAAGAAACCTGAAAATCCAGAGACAAAGCTAGCTATTTCTGTGAACCTCTAGAGATT 7755
Qy 5524 TGGCCCTGCTCTTCAACACAAACAGAGAGAAAGTAACTAAATCATTAATCCCCATG 5583
Db 7756 TGGCCCTGCTCTTCAACACAAACAGAGAGAAAGTAACTAAATCATTAAT - CCCCATG 7814
Qy 5584 GSCCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCTSTTTTACTCTCACTGCA 5643
Db 7815 GCCCTCCCTTATCATATTTTCTCTTATCTGTTTCTTTTACCTCTTCACTCTCACTGCA 7874
Qy 5644 CCCCCCTCATSCCCTGTATGACAGAGTAGCTCCCCCTTACMAAGATTCTATGAGAAATG 5703
Db 7875 CCCCCCTCATSCCCTGTATGACAGAGTAGCTCCCCCTTACMAAGATTCTATGAGAAATG 7934
Qy 5704 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTSTPAAGGAAACCCCAAC 5763
Db 7935 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTCTPAAGGAAACCCCAAC 7994
Qy 5764 TTCACTGCGCCACACCCATATGCCCCGCAACTGCTATCACTTGCCACTTTTGCAATGAT 5823
Db 7995 TTCACTGCGCCACACCCATATGCCCCGCAACTGCTATCACTTGCCACTTTTGCAATGAT 8054
Qy 5824 GCAAAATCTCATTTTGGACAGAGAAATGATTAATCCTAGTTGCTTGAAGAGACTTGA 5883
Db 8055 GCAAAATCTCATTTTGGACAGAGAAATGATTAATCCTAGTTGCTTGAAGAGACTTGA 8114
Qy 5884 GTCACTGCTGTGGACTTACTTCAACCAACCTGATGCTGATGAGGGGTGAGTTCAA 5943
Db 8115 GTCACTGCTGTGGACTTACTTCAACCAACCTGATGCTGATGAGGGGTGAGTTCAA 8174
Qy 5944 GATACGGCAGAGAGAAACATGTAAAGAGTAATCTCCAACTCAACCGGGGTACATGAC 6003
Db 8175 GATACGGCAGAGAGAAACATGTAAAGAGTAATCTCCAACTCAACCGGGGTACATGAC 8234
Qy 6004 ACCCTGAGCCCTCAAAAGACTAGATGCTCAAACTACATGAAACCTCCGATACCAT 6063
Db 8235 ACCCTGAGCCCTCAAAAGACTAGATGCTCAAACTACATGAAACCTCCGATACCAT 8294
Qy 6064 ACTGAGCTGTAGACTTATTAATACACCTCACTGAGGCTCAATGAGGTCTGGCCCA 6123
Db 8295 ACTGAGCTGTAGACTTATTAATACACCTCACTGAGGCTCAATGAGGTCTGGCCCA 8354
Qy 6124 AACCTTACTAAGTGTGATATGCTCCCTGAACTTCAAGCCATATGTTCAATCCCT 6183
Db 8355 AACCTTACTAAGTGTGATATGCTCCCTGAACTTCAAGCCATATGTTCAATCCCT 8414
Qy 6184 GTAACCTAAACAATGGAACAACTTGACAGACAGAAATAAACACCACTCCGTTTATGTA 6243
Db 8415 GTAACCTAAACAATGGAACAACTTGACAGACAGAAATAAACACCACTCCGTTTATGTA 8474
Qy 6244 CCTCTTGTTCCTCAATSTGAAATTAACCATCACTCAAACTCACTGTGTAAATTTAGC 6303
Db 8475 CCTCTTGTTCCTCAATSTGAAATTAACCATCACTCAAACTCACTGTGTAAATTTAGC 8534
Qy 6304 AATACTACATACACAACTCCCAATGATCAAGTGGGTAACTCTCCCAACAATA 6363
Db 8535 AATACTACATACACAACTCCCAATGATCAAGTGGGTAACTCTCCCAACAATA 8594
Qy 6364 GTCTGCTACCCCTCAAGAAATTTTTGTCTGTGTGATCACTCAAGCTATCGTTTGAAT 6423
Db 8595 GTCTGCTACCCCTCAAGAAATTTTTGTCTGTGTGATCACTCAAGCTATCGTTTGAAT 8643
Qy 6424 GGCCTTCAGAAATATATGCTCTCTCATTTTAAAGTCCCCATATGCACTATACACT 6483
Db 8655 GGCCTTCAGAAATATATGCTCTCTCATTTTAAAGTCCCCATATGCACTATACACT 8714
Qy 6484 GAAACAAGATTATACAGTTATGCTATCTAAGCCCGCAACAAAGAGTACCATCTT 6543
Db 8715 GAAACAAGATTATACAGTTATGCTATCTAAGCCCGCAACAAAGAGTACCATCTT 8774
Qy 6544 CCTTTGTATAGAGCAGAGTGTCTAGTGACATAGATCTGCGCATTTGGCGGTATCA 6603
Db 8775 CCTTTGTATAGAGCAGAGTGTCTAGTGACATAGATCTGCGCATTTGGCGGTATCA 8834
Qy 6604 ACTCTTACTAGATTCTATCAAACTATCTCAAGAACTAAATGGGGAACATGGAACGGGTC 6663
Db 8835 ACTCTTACTAGATTCTATCAAACTATCTCAAGAACTAAATGGGGAACATGGAACGGGTC 8894
Qy 6664 GCGGACCTCGGTGACCTTGAAGATCACTTAACCTCCCTAGAGAGATGTCCTTCA 6723
Db 8895 GCGGACCTCGGTGACCTTGAAGATCACTTAACCTCCCTAGAGAGATGTCCTTCA 8954
Qy 6724 AATGAAAGACTTTTAACTGCTTAAACCGCTGABAGAGGGGAAACCTGTTATTTTAAAGG 6783
Db 8955 AATGAAAGACTTTTAACTGCTTAAACCGCTGABAGAGGGGAAACCTGTTATTTTAAAGG 9014
Qy 6784 GAAAGATGCTGTATTTATTTATTAATCAATCCGAAATGCTCACTGAGAAAGTTAAGAAAT 6843
Db 9015 GAAAGATGCTGTATTTATTTATTTAATCAATCCGAAATGCTCACTGAGAAAGTTAAGAAAT 9074
Qy 6844 CSAAGTGAATTAACAAGTAAAGAGAGAGACTTCSAAACACTGAGACCTCGGGGCTCCTC 6903
Db 9075 CSAAGTGAATTAACAAGTAAAGAGAGAGACTTCSAAACACTGAGACCTCGGGGCTCCTC 9134
Qy 6904 AGCCBATGAGATGCGCTGAGATTTCTCCCTCTTGAAGACCTCTAGAGCTATTAATTTGCTA 6963
Db 9135 AGCCBATGAGATGCGCTGAGATTTCTCCCTCTTGAAGACCTCTAGAGCTATTAATTTGCTA 9194
Qy 6964 CTCCTCTTTGGAACCTGATCTTTTAACTCTCTTGTAACTTTGTCTTTCAGAAATGAA 7023
Db 9195 CTCCTCTTTGGAACCTGATCTTTTAACTCTCTTGTAACTTTGTCTTTCAGAAATGAA 9254
Qy 7024 GCTGTAAATTAACAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7083
Db 9255 GCTGTAAATTAACAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9314
Qy 7084 CTGAGACCGGCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7143
Db 9315 CTGAGACCGGCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9374
Qy 7144 GAAATCTAGCTGACACCTCTTACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7203
Db 9375 GAAATCTAGCTGACACCTCTTACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9434
Qy 7204 SGTGGGCAACCTCCCAACAGACCTTAGGTTTCCGTTGAGATGGGGGCTGAGAGAC 7263
Db 9435 SGTGGGCAACCTCCCAACAGACCTTAGGTTTCCGTTGAGATGGGGGCTGAGAGAC 7323
Qy 7264 AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCCTAAGCTAGSTGGGAAAGTAC 7323
Db 9495 AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCCTAAGCTAGSTGGGAAAGTAC 9554
Qy 7324 ACATCAACTTTTAAACAGGGGCTTGCACTTAGTCAACCTGACCAATGAGAGACTC 7383
Db 9555 ACATCAACTTTTAAACAGGGGCTTGCACTTAGTCAACCTGACCAATGAGAGACTC 9614
Qy 7384 ACTAAATGCTATTTAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7443
Db 9615 ACTAAATGCTATTTAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9674
Qy 7444 AGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7503
Db 9675 AGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9734
Qy 7504 CAAACCCCTTTGGGCTCCCTCCCTTTGTATGAGAGCTTGTGTTTCAATGCTATTTCACT 7563
Db 9735 CAAACCCCTTTGGGCTCCCTCCCTTTGTATGAGAGCTTGTGTTTCAATGCTATTTCACT 9794
Qy 7564 ATTAAATCTTGACATGCTGCR 7582
Db 9795 ATTAAATCTTGACATGCTGCA 9813
```

```

RESULT 7
AX329572 . 56093 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 81 from Patent WO0194629.
DEFINITION AX329572
ACCESSION AX329572
VERSION AX329572.1 GI:18102550
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Hortigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 81.13-DEC-2001.
JOURNAL Avalon Pharmaceuticals (US)
FEATURES
Source 1..56093
Location/Qualifiers
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 85.0%; Score 6448; DB 6; Length 56093;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;
QY 606 GATGGGAAACGTTCCCGCAGACAAAACGCCCTTAAGAGCTATTTGGAAAAATTGGGA 665
D 30953 GATGGGAAACGTTCCCGCAGACAAAACGCCCTTAAGAGCTATTTGGAAAAATTGGGA 31012
QY 666 MCAATTTGACCCCTGACACTAAGAAAAGAAAGCACTATCTCTGAGGCGCGCTG 725
D 31013 CCAATTTGACCCCTGACACTAAGAAAAGAAAGCACTATCTCTGAGGCGCGCTG 31072
QY 726 GCACTCTGAGGAGAGTAAATTATTAACACCATCTTACAGCTAGACCTTTTGTAGA 785
D 31073 GCATCTCTGAGGAGAGTAAATTATTAACACCATCTTACAGCTAGACCTTTTGTAGA 31132
QY 786 AAGGCAATGGAAGTGAAGTGCATTAAGCAAACTTTCTTTCAATTAAGAGCAACTCAG 845
D 31133 AAGGCAATGGAAGTGAAGTGCATTAAGCAAACTTTCTTTCAATTAAGAGCAACTCAG 31192
QY 846 AATTATGTAAGAAAGTGAATTATGCCCCAGAGAAAGCTTCAAGAGTCACTCCCTATC 905
D 31193 AATTATGTAAGAAAGTGAATTATGCCCCAGAGAAAGCTTCAAGAGTCACTCCCTATC 31252
QY 906 CCAGCATCCCGCACTCTTCCCCAMTATTAAGAACCCCTTCAACCCAAATGCTCCAA 965
D 31253 CCAGCATCCCGCACTCTTCCCCAMTATTAAGAACCCCTTCAACCCAAATGCTCCAA 31312
QY 966 AAGGAGATGACAAAAGGTTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTATGA 1025
D 31313 AAGGAGATGACAAAAGGTTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTATGA 31372
QY 1026 CCCCCCAAGCAGGAGGAGAAAGAAATTGGCCAGCAGAGAGCATGTCCTTTTXX 1085
D 31373 CCCCCCAAGCAGGAGGAGAAAGAAATTGGCCAGCAGAGAGCATGTCCTTTTXX 31431
QY 1086 TCTCCAGACTTAAAGCAAAATTAAGACAGACTTAGTAATTTCTAGATTAATCTGATGG 1145
D 31432 TCTCCAGACTTAAAGCAAAATTAAGACAGACTTAGTAATTTCTAGATTAATCTGATGG 31491
QY 1146 CTATATTTGTTTAAAGGTTAGAGCAATTTCTTGAATCTGACATGAGAGATATATA 1205
D 31492 CTATATTTGTTTAAAGGTTAGAGCAATTTCTTGAATCTGACATGAGAGATATATA-A 31550
QY 1206 TGTGCTGCTAAATAGACACTTAACCCCAATGAGAGAGTGCACCATTAATCTGAGGCT 1265
D 31551 TGTGCTGCTAAATAGACACTTAACCCCAATGAGAGAGTGCACCATTAATCTGAGGCT 31610

```

```

QY 1266 GAGGTTTGGGATCTCTGTTATCTCACTCAAGTCAATGATTAAGATGACAAACGAGG 1325
D 31611 GAGGTTTGGGATCTCTGTTATCTCACTCAAGTCAATGATTAAGATGACAAACGAGG 31667
QY 1326 AAAGANATGATTTCCCAACAGGCGCAGCAGGAGTCCCAAGTCTASACCTCATTTGGGAGC 1385
D 31668 AAAGANATGATTTCCCAACAGGCGCAGCAGGAGTCCCAAGTCTASACCTCATTTGGG 31724
QY 1386 ACAGAAATCAATGATGAGGAGTGTGCTGACAGCAATTTGCTAATCTTGTGTCTAATA 1445
D 31725 ACAGAGATCAAGACAT -GAGATTTGGTGTGACAGCAATTTGCTAATCTTGTGTCTAATA 31783
QY 1446 GGACTAAGGAAATCTAAGAGAAATCTATATTAATTAATGATGATGATCCCATTAACA 1505
D 31784 GGACTAAGGAAATCTAAGAGAAATGAGTCTATTAATTAATTAATGATGATGATCCCATTAACA 31842
QY 1506 GGGGAGGAAAGAAATCTCTGCTCTTCTGAGAGACTAAGGAGGATTTGAGGAGG 1565
D 31843 -GGGAGGAAAGAAATCTCTGCTCTTCTGAGAGACTAAGGAGGATTTGAGGAGG 31901
QY 1566 GTGCTCTCTGTCACTGACTCTTCTGAGAGGCAACTAATCTTAAGCGTAAGTTATCA 1625
D 31902 GTGCTCTCTGTCACTGACTCTTCTGAGAGGCAACTAATCTTAAGCGTAAGTTATCA 31961
QY 1626 CTGAGTCACTGAGCACTTAG -AAAAAATCTTCAAAAGTCTGCGCTAGCCCGAGCAAA 1684
D 31962 CTGAGTCACTGAGCACTTAGAAAAAATCTTCAAAAGTCTGCGCTAGCCCGAGCAAA 32021
QY 1685 ACTTAGAAACCCCTATTTGAATTTGCACTGCGGTTTATTAATTAAGATCAGAGAGAGC 1744
D 32022 ACTTAGAAACCCCTATTTGAATTTGCACTGCGGTTTATTAATTAAGATCAGAGAGAGC 32081
QY 1745 AAGCGAAACAGACAAACGGGATTTAAAAAAGGCCACCGCTTATGTCATGACCTCAGG 1804
D 32082 AAGCGAAACAGACAAACGGGATTTAAAAAAGGCCACCGCTTATGTCATGACCTCAGG 32141
QY 1805 CAATGATCTTTGAGAGTCTGAGAAAGGAAAGGAAAGCTGGGCAATTTGAATGCTTAATAG 1864
D 32142 CAATGATCTTTGAGAGTCTGAGAAAGGAAAGGAAAGCTGGGCAATTTGAATGCTTAATAG 32201
QY 1865 GCTTGCTTCCAGTCCGAGCTCAAGAGACACTTTAAAAAAGATTGCCAAGTGAAGTAAG 1924
D 32202 GCTTGCTTCCAGTCCGAGCTCAAGAGACACTTTAAAAAAGATTGCCAAGTGAAGTAAG 32261
QY 1925 CCGCCCTTCTGTCATGCGCCCTTATTTCAAGGAAATCACTGAAAGGCCCACTGCCCAAG 1984
D 32262 CCGCCCTTCTGTCATGCGCCCTTATTTCAAGGAAATCACTGAAAGGCCCACTGCCCAAG 32321
QY 1985 GGAACAAAGTCTTTTGAAGTCAAGAACCACTTAACAGATGATTCAGACAGACAGTGAAGG 2044
D 32322 GGAACAAAGTCTTTTGAAGTCAAGAACCACTTAACAGATGATTCAGACAGACAGTGAAGG 32381
QY 2045 TGCCCTGGGGGAGGAGGCAATCCCATGCACTACCCCTCAAGAGCCCTGGGATATGCTTGAAC 2104
D 32382 TGCCCTGGGGGAGGAGGCAATCCCATGCACTACCCCTCAAGAGCCCTGGGATATGCTTGAAC 32441
QY 2105 ATTGAGGCGCAGAAAGTTGTCTCTGAGCACTGGTGGGCTTTCTTATGCTTACTCTTC 2164
D 32442 ATTGAGGCGCAGG -AGGTTGTCTCTGAGCACTGGTGGGCTTTCTTATGCTTACTCTTC 32500
QY 2165 TGTCCCGGACAACTGTCTCTCAAGTCTGTCACTATTCTGAGAGGGGTCCNTAAGAGGGCA 2224
D 32501 TGTCCCGGACAACTGTCTCTCAAGTCTGTCACTATTCTGAGAGGGGTCC -TAAAGAGGGCA 32558
QY 2225 GTCACTAGATCTTTTCCGAGCACTTAAGTATTAAGCTGGGAGCTTATTTCTTTTAC 2284
D 32559 GTCACTAGATCTTTTCCGAGCACTTAAGTATTAAGCTGGGAGCTTATTTCTTTTAC 32616
QY 2285 ATGCTTTTCTAATTAATGTTGAAAGCCCACTACTCTGTTAAGGAGAGCAATTTAGCA 2344
D 32617 ATGCTTTTCTAATTAATGTTGAAAGCCCACTACTCTGTTAAGGAGAGCAATTTAGCA 32676

```

```
QY 2345 AAGAGGGGCAATTATACCTGAACATAGAGAAAGAAACACCCGTTGTGTTCCCTG 2404
DB 32677 AAGCAGGGGGCAATTATACACTGAAACATAGAGAAAGAAACACCCGTTGTGTTG - CCCCG 32735
QY 2405 CTTGAGGAAGGAATTAATCTCTGAAGCTGCGGCAAGAAAGACAAATATGAGAGCAAA 2464
DB 32736 CTTGAGGAAGGAATTAATCTCTGAAGCTGCGGCAAGAAAGACAAATATGAGAG - CAAA 32794
QY 2465 GAATGCGCGTCTGTTCAAGTTAACTAAAGATTCCACTTCCTTCCCTACCAAGCA 2524
DB 32795 GAATGCGCGTCTGTTCAAGTTAACTAAAGATTCCACTTCCTTCCCTACCAAGCA 32854
QY 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAAGATTGTTAAGACTTAAAGC 2584
DB 32855 GTACCCCTCAGACCCCAAGGCCCAACAGGACTCCAAAAGATTGTTAAGACTTAAAGC 32914
QY 2585 CCAAGGCTTAGTAAACCATGATTAACCTCCCTGAGTAATCCGATAGTGGATTGAGAGG 2644
DB 32915 CCAAGGCTTAGTAAACCATGATTAACCTCCCTGAGTACTCC-----AATTTTAGAG 32967
QY 2645 CACAGAAACCCAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAATCAATGAGGC 2704
DB 32968 TACAGAAACCCAGACAGACAGTGA - GATTAGTGAAGATCTCAGGATTAATCAAT - GAGGC 33025
QY 2705 GGTGTCTTTTATACCAAGCTGTACTAGCCCTTACTGTGMYTCCCAATACAGA 2764
DB 33026 TGTGTCTCTATAGCCAGCTGTACTAGCCCTTACTGTGCTTCCCAATACAGA 33085
QY 2765 GGAAGCAGAGTGTATTAACSTCTGAGACTTMAAGATGCTTCTGATCCCTGTACA 2824
DB 33086 GGAAGCAGAGTGTATTAACSTCTGAGACTTMAAGATGCTTCTGATCCCTGTACA 33145
QY 2825 TCCTGACTCTCAATCTTGTTGCTTGAAGATCTTCAACCCAGCATCTCACTAC 2884
DB 33146 TCCTGACTCTCAATCTTGTTGCTTGAAGATCTTCAACCCAGCATCTCACTAC 33205
QY 2885 CTGACATRTTTTACCCCAAGGGTTCAAGGATAGYCCCATCTATTGTCAGGCACTTAGC 2944
DB 33206 CTGACATRTTTTACCCCAAGGGTTCAAGGATAGTCCCATCTATTGTCAGGCACTTAGC 33265
QY 2945 CCAAGACTTGAAGCAATCTTACTGAGACCTTTGCTCTGKTAAGTGAATTT 3004
DB 33266 CCAAGACTTGAAGCAATCTTACTGAGAC - CTGTGCTCTGCTGATGAGTGAATTT 33323
QY 3005 ACTTTGCGCYGCTRTTGAAGAACTTGSCATCAAGCCACCCAGRCCTTMAATTT 3064
DB 33324 ACTTTGCGCGCCATCAAGAACTTGSCATCAAGCCACCCAGRCCTTCAATTT 33383
QY 3065 CCTGCGYACCTGTGCTACAMGTTTCCAAACSAARGCTCACTCTGCTACAGCAGT 3124
DB 33384 CCTGCGCTACCTGTGCTACAMGTTTCCAAACSAARGCTCACTCTGCTACAGCAGT 33443
QY 3125 TAAATTACTTGAAGTAAATTTTCCAAAGGCAACARGGCCTCAGTGAAGAAATTTCA 3184
DB 33444 ----TACTTGAAGGCTAAATTTTCCAAAGGCAACARGGCCTCAGTGAAGAACTCA 33499
QY 3185 GCCTATACGCTTATCTCATTCYCAAAACCTTAAAGCACTAAGRRTCTTGCGRT 3244
DB 33500 GCCTATACGCTTATCTCATTCYCAAAACCTTAAAGCACTAAGRRTCTTGCGGT 33559
QY 3245 AAAYAGYTTTCGCGAAMATGATTTCCAGATWTGCGRAAATAGCCAGYCATTAATA 3304
DB 33560 AATAGGTTTTCGCGAAMATGATTT - CCAAGATWTGCGRAAATAGCCAGYCATTAATA 33618
QY 3305 CAAATTAATTAAGAAACTCAGAAAGCAATACCATTTATATAGATGAYAMCTGAGYMR 3364
DB 33619 CAAATTAATTAAGAAACTCAGAAAGCAATACCATTTATATAGATGAGYAMCTGAGYMR 33678
QY 3365 AAGTGGCTTTCAGAGCCCTTAAAGAGGCTTAAACCAAGYCCAGGTTAAAGTTGCC 3424
DB 33679 AAGTGGCTTTCAGAGCC-----TAAACCAAGCCCAAGGTTAAAGTTGCC 33724
QY 3425 AACRGGCGAAGACTTTTSTYATATRTCAAGAAAAAACAGRAAYAGCTCTRGAGTCC 3484
DB 33725 AACRGGCGAAGACTTTTCTATATRTCAAG - AAAAAACAGAAATAGCTCTAGAGTCC 33783
QY 3485 TTACACAGRTCCRAGGAYAGCTTTGCAACCTRTGCGYACTGASTAAAGAAATGATG 3544
DB 33784 TTACACAGATCCGAGGAYAGCTTTGCAACCTRTGCGYACTGASTAAAGAAATGATG 33843
QY 3545 TAGTGGCAAGGGTGTGCTGATTTTAYGGGTGTGCTGAGTACAGYKTAAGT 3604
DB 33844 TAGTGGCAAGGGTGTGCTGATTTTAYGGGTGTGCTGAGTACAGYKTAAGT 33903
QY 3605 CTGAAGCATTAAATTAATACAGGAGAGATCTTACTGTGTGACATCTCATGAKGTGA 3664
DB 33904 CTGAAGCATTAAATTAATACAGGAGAGATCTTACTGTGTGACATCTCATGAKGTGA 33963
QY 3665 AYRGCTACTCACTGTCTAAAGAGACTTGTGCTGTGACACACCTGTTTACTTAAATRTG 3724
DB 33964 AYRGCTACTCACTGTCTAAAGAGACTTGTGCTGTGACACACCTGTTTACTTAAATRTG 34023
QY 3725 AGGCTCTATTAGTGAAGGCGAGGCTGACCTGTGCACTTGTCACCTTAAACCGAG 3784
DB 34024 AGGCTCTATTAGTGAAGGCGAGGCTGACCTGTGCACTTGTCACCTTAAACCGAG 34083
QY 3785 YCNCAATTTCTTCAAGACATGAGAAAGATARAAYATPACTGTCAACAAATATTTCTC 3844
DB 34084 CCACATTTCTTCAAGACATGAGAAAGATARAAYATPACTGTCAACAAATATTTCTC 34143
QY 3845 AAACCTATGCACTGAGGGGACTTGTAGAGTTCCYTTGACTGATCCYACCTTCAAC 3904
DB 34144 AAACCTATGCACTGAGGGGACTTGTAGAGTTCCYTTGACTGATCCCAAC - TCAAC 34202
QY 3905 TTGTATCTGATGGAAGTCTCTTGTGAAAAAGAAAGCTGAAAAAGGGGTATAGCACTG 3964
DB 34203 TTGTATCTGATGGAAGTCTCTTGTGAAAAAGAAAGCTGAAAAAGGGGTATAGCACTG 34262
QY 3965 GTCAGATTAATGGAATATTTGAAAGTATCCCTCACTCCAGAACTAGTCTYAGCTR 4024
DB 34263 GTCAGATTAATGGAATATTTGAAAGTATCCCTCACTCCAGAACTAGTCTYAGCTR 34332
QY 4025 GCAGAACTAATAGCTTCATYTKGGGCACTAGAAATTAAGAGAAAGGAAATATTA 4084
DB 34323 GCAGAACTAATAGCTTCATYTKGGGCACTAGAAATTAAGAGAAAGGAAATATTA 34382
QY 4085 TATACAGCTCTRTATGCTYATCTGATGTCATGCGCATGMRGAATATGSRABAGA 4144
DB 34383 TATACAGCTCTRTATGCTYATCTGATGTCATGCGCATGMRGAATATGSRABAGA 34442
QY 4145 AGGGAATTCCTAATCTTGYAGAGAACCTATCAACATCAGAAAGCCATTAGARATTA 4204
DB 34443 AGGGAATTCCTAATCTTGYAGAGAACCTATCAACATCAGAAAGCCATTAGARATTA 34502
QY 4205 TTATYGGCKGTACAGAAACCTATAGAGGTGMACTTTACCTGCTGCGGTTCATCANAA 4264
DB 34503 TTATYGGCKGTACAGAAACCTATAGAGGTGMACTTTACCTGCTGCGGTTCATCANAA 34562
QY 4265 GGAAGAAAGGGGAAATPASAAGAAATGACCAAGCAATATTTGAAGCAAAAGGCTGCA 4324
DB 34563 GGAAGAAAGGGGAAATPASAAGAAATGACCAAGCAATATTTGAAGCAAAAGGCTGCA 34622
QY 4325 AGGCAAGACCTCATTAAGAAATGCTTATTAATCTTCCCTTATAGATAGGATCCCTTC 4384
DB 34623 AGGCAAGACCTCATTAAGAAATGCTTATTAATCTTCCCTTATAGATAGGATCCCTTC 34682
QY 4385 CGGGAACCAAGCCCACTACTCAGCAGAGAAACAGAAATGGGGAACCTCAGAGG - CAG 4443
DB 34683 CGGGAACCAAGCCCACTACTCAGCAGAGAAACAGAAATGGGGAACCTCAGAGG - CAG 34742
QY 4444 TTTTCTCCCTCGGAGAGGTAGCACTGTAAGAAAGGAAATATTTTGTGCTGCACTAT 4503
DB 34743 TTTTCTCCCTCGGAGAGGTAGCACTGTAAGAAAGGAAATATTTTGTGCTGCACTAT 34802
QY 4504 CCAATGGAATTAATTAACCTTCACTCAACCTTCACTTATAGCATGATACCACTCA 4563
```

Db 34803 CCAATGGAATTAATAAACCCTTCATCAAAACCTTCACTTAGACATCATGACACCA 34862
Qy 4564 TCATATGGCCAAATCATATTTACTGACACAGCCCTTTCAAACCTATCAAGCABATAT 4623
Db 34863 TCAGATGGCCAAATCATATTTACTGACACAGCCCTTTCAAACCTATCAAGCABATAT 34922
Qy 4624 CAGGCGCTGTGAATGTGGCCABARAATAATCCCTGCTATATGCGCAAGCTCTTCAAG 4683
Db 34923 CAGGCGCTGTGAATGTGGCCABARAATAATCCCTGCTATATGCGCAAGCTCTTCAAG 34982
Qy 4684 ARAACAABAAACAGGCCATTAACCTGTABARAATCTGGCACTGATTTTACCAAGAGCC 4743
Db 34983 AGAACAABAAACAGGCCATTAACCTGTABARAATCTGGCACTGATTTTACCAAGAGCC 35042
Qy 4744 AAACCTCAGGATTTCAAGTATCTAAGTCTGGGATGATTTTCAAGGCTTGGGCAAG 35102
Db 35043 AAACCTCAGGATTTCAAGTATCTAAGTCTGGGATGATTTTCAAGGCTTGGGCAAG 35102
Qy 4804 GCCTTCCCTGTAGACAGAAAGGCCAAGGTAATAAGGCACTAGTTCATGAATA 4863
Db 35103 GCCTTCCCTGTAGACAGAAAGGCCAAGGTAATAAGGCACTAGTTCATGAATA 35162
Qy 4864 ATTCCAGATTCGAGCTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTCCAGGCC 4923
Db 35163 ATTCCAGATTCGAGCTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTCCAGGCC 35222
Qy 4924 ACAGTAACCAAGGAGTATCCAGAGCGTTAGATATACGATATCACTTACCTGCGCTGA 4983
Db 35223 ACAGTAACCAAGGAGTATCCAGAGCGTTAGATATACGATATCACTTACCTGCGCTGA 35282
Qy 4984 AGGCCAAGTCTTCAAGGAGTGAAGAAATGAAATGAATACTCAAGGCACTCTAAA 5043
Db 35283 AGGCCAAGTCTTCAAGGAGTGAAGAAATGAAATGAATACTCAAGGCACTCTAAA 35342
Qy 5044 AAGCAAAACCGAGAAACCCACCTCAATGCGCTGTGCTATAGCCTTAAAAAGA 5103
Db 35343 AAGCAAAACCGAGAAACCCACCTCAATGCGCTGTGCTATAGCCTTAAAAAGA 35402
Qy 5104 ATCTGCAACTTTCCCAAAAAGCAGAGACTTAAAGCCATAGCAATGCTGTATGAAAGGCC 5163
Db 35403 ATCTGCAACTTTCCCAAAAAGCAGAGACTTAAAGCCATAGCAATGCTGTATGAAAGGCC 35462
Qy 5164 TTCAATAACAATGACCTTGTGCTTGAACCAAGACAGCCAACTTATGTCAGACATCACT 5223
Db 35463 TTCAATAACAATGACCTTGTGCTTGAACCAAGACAGCCAACTTATGTCAGACATCACT 35522
Qy 5224 CCTTAGCCAAATATCAACAAGTCTTAAAACTTAAAGAAACCTATCCCTGAGAAAGG 5283
Db 35523 CCTTAGCCAAATATCAACAAGTCTTAAAACTTAAAGAAACCTATCCCTGAGAAAGG 35582
Qy 5284 GAAAGAAGACTATTCACACCCMWGTGACATGTATTAATGCAAGTCCCTTCTCTAATTC 5343
Db 35583 GAAAGAAGACTATTCACACCCMWGTGACATGTATTAATGCAAGTCCCTTCTCTAATTC 35642
Qy 5344 CATCCCTAGATACATCTGAGGAAAGACCTTACCCAGTCAATTTTATCTACCCCACTGG 35702
Db 35643 CATCCCTAGATACATCTGAGGAAAGACCTTACCCAGTCAATTTTATCTACCCCACTGG 35702
Qy 5404 TTAAAGTGGCTGAGTGGAGTCTTGATATCATCACTTGAATCAAACTCTGATATGCG 5463
Db 35703 TTAAAGTGGCTGAGTGGAGTCTTGATATCATCACTTGAATCAAACTCTGATATGCG 35762
Qy 5464 CAAAGGAACCTGAAAAATCAGAGAGCAACGCTAGTATCTCTGAAACCTCTAGAGGAT 5523
Db 35763 CAAAGGAACCTGAAAAATCAGAGAGCAACGCTAGTATCTCTGAAACCTCTAGAGGAT 35822
Qy 5524 TGGCGCTGCTCTTCAAACAACAAGAGGAAAGTAACTAAATCAATAAATCCGCCATG 5583
Db 35823 TGGCGCTGCTCTTCAAACAACAAGAGGAAAGTAACTAAATCAATAAAT-CCGCCATG 35881
Qy 5584 GSCCTCCCTATCATATTTTCTCTKTATGTTTATCTTTCCTTTTCACTCTCACTGA 5643
Db 35882 GSCCTCCCTATCATATTTTCTCTKTATGTTTATCTTTCCTTTTCACTCTCACTGA 35941

Qy 5644 CCCCTCAATGCCGCTGTATGACCAAGTATCCCTTACCMAGATTTCTATGAGATG 5703
Db 35942 CCCCTCAATGCCGCTGTATGACCAAGTATCCCTTACCMAGATTTCTATGAGATG 36001
Qy 5704 CAGCGTCCCGAAATATTTGATGCGCCATGTATAGAGTCTTTSTAAGGAACCCCAAC 5763
Db 36002 CAGCGTCCCGAAATATTTGATGCGCCATGTATAGAGTCTTTSTAAGGAACCCCAAC 36061
Qy 5764 TTCACTGCGCAACCCATATGCCCCGCAATGCTATACATGCGCACTCTTTGATGAT 5823
Db 36062 TTCACTGCGCAACCCATATGCCCCGCAATGCTATACATGCGCACTCTTTGATGAT 36121
Qy 5824 GCAATATCTATATTGAGACAGAAAAATGATTAATCTATGTTGCTCTGAGAGACTTGA 5883
Db 36122 GCAATATCTATATTGAGACAGAAAAATGATTAATCTATGTTGCTCTGAGAGACTTGA 36181
Qy 5884 GTCATGTCTGTGGACTTACCTTCAACCAATGATATGTTGATGAGGGGTGAGTTCAA 5943
Db 36182 GTCATGTCTGTGGACTTACCTTCAACCAATGATATGTTGATGAGGGGTGAGTTCAA 36241
Qy 5944 GATCAGGCAAGAAAAACATTTAAAGAAATATCTCCCACTCACCGGGTATGAGGC 6003
Db 36242 GATCAGGCAAGAAAAACATTTAAAGAAATATCTCCCACTCACCGGGTATGAGGC 36301
Qy 6004 ACCCTAGCCCTTACAAGGACTAGATCTTCAAAACTACATGAAACCTTCCGTATCCAT 6063
Db 36302 ACCCTAGCCCTTACAAGGACTAGATCTTCAAAACTACATGAAACCTTCCGTATCCAT 36361
Qy 6064 ACTGCGCTGTGAAGCTTATTTAAATACCACTTCACTGAGGCTCCATGAGTCTCGGCCAA 6123
Db 36362 ACTGCGCTGTGAAGCTTATTTAAATACCACTTCACTGAGGCTCCATGAGTCTCGGCCAA 36421
Qy 6124 AACCTTACTAATCTGTGGAATATGCTCCCTGAACTTCAAGGCAATGTTTCAATCCCT 6183
Db 36422 AACCTTACTAATCTGTGGAATATGCTCCCTGAACTTCAAGGCAATGTTTCAATCCCT 36481
Qy 6184 GTACCTGAACATGGAACAACCTTACAGCAGAAATTAACAACCTTCCGTTTATGAGA 6243
Db 36482 GTACCTGAACATGGAACAACCTTACAGCAGAAATTAACAACCTTCCGTTTATGAGA 36541
Qy 6244 CCTCTGTGTTCAATSTGAAATTAACCATPACTCAAACTCACTGCTGTATTAATGAC 6303
Db 36542 CCTCTGTGTTCAATSTGAAATTAACCATPACTCAAACTCACTGCTGTATTAATGAC 36601
Qy 6304 AATATCATPACACACCAACTCCATATGATCATGAGTGGGTAATCTCTCCCAACAATA 6363
Db 36602 AATATCATPACACACCAACTCCATATGATCATGAGTGGGTAATCTCTCCCAACAATA 36661
Qy 6364 GTCTGCTACCCCTAGAAATATTTTGTCTGAGGTAACCTCAGCCTATCGTTTGAAT 6423
Db 36662 GTCTGCTACCCCTAGAAATATTTTGTCTGAGGTAACCTCAGCCTATCGTTTGAAT 36721
Qy 6424 GGCCTTCAAGATCTATGCTTCTCTCATTTCTTATGAGCCCTYATGCACTTCACT 6483
Db 36722 GGCCTTCAAGATCTATGCTTCTCTCATTTCTTATGAGCCCTATGCACTTCACT 36781
Qy 6484 GAACAAATTTATACATGATATGATATCTTAAGCCCGCAACAAGAGTACCCATCTCT 6543
Db 36782 GAACAAATTTATACATGATATGATATCTTAAGCCCGCAACAAGAGTACCCATCTCT 36841
Qy 6544 CCTTTGTATAGAGACAGAGTCTAGTGTCACTAGGTACTGTGCACTTGGCGGTATCA 6603
Db 36842 CCTTTGTATAGAGACAGAGTCTAGTGTCACTAGGTACTGTGCACTTGGCGGTATCA 36901
Qy 6604 ACCTTACTAGTCTTCTATCAAACTATCTCAAGACTAATTTGGGACATGGAACGGGTC 6663
Db 36902 ACCTTACTAGTCTTCTATCAAACTATCTCAAGACTAATTTGGGACATGGAACGGGTC 36961
Qy 6664 GCCGACTCCCTGTGACCTTTCAGAGATCACTTATCTCCCTTACAGAGAGTATGCTCTG 6723
Db 36962 GCCGACTCCCTGTGACCTTTCAGAGATCACTTATCTCCCTTACAGAGAGTATGCTCTG 37021

QY	6724	AATGGAAGAGCTTTAGACATCTGCTAAACCGCTGACAGAGGGGAAACCTGTTATTTTAAAGG	6783		
Db	37022	AATGGAAGAGCTTTAGACATCTGCTAAACCGCTGAAAGGGGGAAACCTGTTATTTTAAAGG	37083		
QY	6784	GAAGAATCTGTTATTTATGTTAATCATC	6843		
Db	37082	GAAGAATGCTGTTATTTATGTTAATCATC	37144		
QY	6844	CSAGATGGAATPACAACGTAACAGACAGAGCTTCCAAACACTGGACCTCTGGGCTCTCTC	6903		
Db	37142	CSAGATGGAATPACAACGTAACAGAGAGCTTCCAAACACTGGACCTCTGGGCTCTCTC	37203		
QY	6904	AGCCATGATGATGACCTGATATCTCCCTCTTAAAGACCTCTGACGCTAATTAATTTGCA	6963		
Db	37202	AGCCATGATGATGACCTGATATCTCCCTCTTAAAGACCTCTGACGCTAATTAATTTGCA	37263		
QY	6964	CTCCTCTTTGGACCCCTGATCTTATCTTACCTCTCTTAACTTTGCTCTTCCAGAAATCGAA	7023		
Db	37262	CTCCTCTTTGGACCCCTGATCTTATCTTACCTCTCTTAACTTTGCTCTTCCAGAAATCGAA	37322		
QY	7024	GCTGTAAATCTAAATGAGACCCAAAGATGCAAGTCCAAAGCTAATTAATCCGACAGACC	7083		
Db	37322	GCTGTAAATCTAAATGAGACCCAAAGATGCAAGTCCAAAGCTAATTAATCCGACAGACC	37383		
QY	7084	CTGACCGGGCTGTATGACCCAGATCTGATGTTATGACATTAAGAGCAACCCCTCTGAG	7143		
Db	37382	CTGACCGGGCTGTATGACCCAGATCTGATGTTATGACATTAAGAGCAACCCCTCTGAG	37443		
QY	7144	GAAATCTCAGCTGACAAACCTCTAATGAGCCCAATTAAGAGAGAGAGCTTAAAGCGGT	7203		
Db	37442	GAAATCTCAGCTGACAAACCTCTAATGAGCCCAATTAAGAGAGAGAGCTTAAAGCGGT	37503		
QY	7204	SGTCGGCCAACTCTCCCAACGACCTTAGGTTTCTGTGAGATGGGGACCTGAGAGAC	7263		
Db	37502	SGTCGGCCAACTCTCCCAACGACCTTAGGTTTCTGTGAGATGGGGACCTGAGAGAC	37563		
QY	7264	AGGACTAGCTGAGATTTCTTAGGCTGATTAAGAAATCTTAAGCTTAAAGAGGTGAC	7323		
Db	37562	AGGACTAGCTGAGATTTCTTAGGCTGATTAAGAAATCTTAAGCTTAAAGAGGTGAC	37623		
QY	7324	ACATTCACCTTTAAACAGGGGGCTTGCAACTAGTACACCTGACCAATCAGAGAGCTC	7383		
Db	37622	ACATTCACCTTTAAACAGGGGGCTTGCAACTAGTACACCTGACCAATCAGAGAGCTC	37683		
QY	7384	ACTAAATGCTAATTTAGGCAAAAGACAGAGGTAAAGAAATGCCAAATCATTTATGCTMG	7443		
Db	37682	ACTAAATGCTAATTTAGGCAAAAGACAGAGGTAAAGAAATGCCAAATCATTTATGCTMG	37743		
QY	7444	ACAGCACAGACAGAGAGACATATATGGGATATTAACCCAAATGTTGAGCCGGCAACGG	7503		
Db	37742	ACAGCACAGACAGAGAGAGACATATATGGGATATTAACCCAAATGTTGAGCCGGCAACGG	37803		
QY	7504	CAACCCCTTTGGGATCCCTCCCTCTTAAATGGGAGCTCTGTTTCAATGCTAATTCACCT	7563		
Db	37802	CAACCCCTTTGGGATCCCTCCCTCTTAAATGGGAGCTCTGTTTCAATGCTAATTCACCT	37863		
QY	7564	ATTAAATCTTGACATGCR 7582			
Db	37862	ATTAAATCTTGACATGCA 37880			
RESULT 8	HSAC000064	56093 bp	DNA	linear	PRI 13-NOV-1996
LOCUS	DEFINITION	Human BAC clone R0803M05 from 7q21-7q22, complete sequence.			
ACCESSION	AC000064				
VERSION	AC000064.1	GI:1669369			
KEYWORDS	HTG.				
SOURCE	Organism	Human sapiens (human)			
REFERENCE	1	(Baees 1 to 56093)			

AUTHORS	Pauley, A.
TITLE	The sequence of H. sapiens BAC clone RG083M05
JOURNAL	Unpublished (1996)
REFERENCE	2 (bases 1 to 56093)
AUTHORS	Waterston, R.
TITLE	Direct Submision
JOURNAL	Submitted (13-NOV-1996)
COMMENT	Genome Sequencing Center

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:
This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 9788K. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
VECTOR: pBelo
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H_RG083M05; actual end is at 560933 of H_RG083M05

Practitioner

Source

This clone contains STS SWSS1725.

repeat_region

ग्राम

CDS

```

/gene="WUGSC:H_RG083M05.1"
join(<1360. .1503,4181. .4370,4587. .4774,6422. .6556,

```

9483.	.9547,11631.	.11773,11864.	.12021,13131.	.13296,
14885.	.14988,16349.	.16546,16837.	.16971)	

```
/gene="WUGSC:H_RS083M05.1"  
/note="ATPase; strong similarity to peroxisome  
biogenesis factor 1 (nucleoside diphosphate-  
binding domain)"
```

biosynthesis protein PAS1 (P1D:gil1/2019); coded for by
human CDNA C04279 (NID:gil167530) "
/codon_start=1

```

/coach_state=1
/protein_id="AAB46346.1"
/db_xref="GI:1669371"

```

```
translation="KRLNIQKTEVAFSEAVVMQPSVLLDDLLIAGLPAPV  
/SPDAEORCEILCNVINKKLDCKINKFTDLLQHVAKETGCFVARDFTVLVDRAI
```

SRQSI~~STRE~~KVL~~TT~~LDFO~~KAL~~RGL~~PAS~~LR~~SVN~~LHK~~PRD~~LGD~~NDK~~IGGL~~HEV~~RQ
TIQLPA~~VCL~~KEKE~~Y~~PEL~~FAN~~LPIRQ~~TGI~~LLYGP~~PTG~~KTLLAG~~VI~~ARESR~~MN~~

LTQLDVGEGLOGVYLAATSRPDLIPALLRGRLDKCVYCEPPQVLTISYLE
KGPBLLSKYIGASEQVRDIFIRAGAAKPCILFEDEFESIABRGHDNTGVTDR

QUMTSF LVS RUI LNV LMS DYL PLADVDLQHVAS IDS F IGADLEKALL INAOLE
MULSKMSEILPDESKFNRYLFGSS YSESELGNGTSSD LSSQCLSPSSMTODL
CKDOLFSOBPVV PTAOEGGCOEI TOFOOPDLOPADI SI IKGYPBOSCEGSESMNO

repu

region
complement(4948.5130)
KTRLAISQSHLMTALGHTRPSISEDWKNFAEL"

```
/rpc_family="ALU"
```

```

repeat_region complement(6581. .7133)
                /rpt_family="L1"
repeat_region complement(7767. .8037)
                /rpt_family="ALU"
repeat_region complement(8186. .8472)
                /rpt_family="ALU"
misc_feature 8473. .8625
                /gene="WUGSC:H.RG083M05.1"
                /note="match to human 3' EST H75782 (NID:g1049794), bases
                287-444"
misc_feature 8841. .9161
                /gene="WUGSC:H.RG083M05.1"
                /note="match to human 5' EST H75921 (NID:g1050050), bases
                21-348"
misc_feature 9481. .9547
                /gene="WUGSC:H.RG083M05.1"
                /note="match to human 5' EST N2627 (NID:g1130501), bases
                276-343"
repeat_region complement(12612. .12907)
                /rpt_family="ALU"
misc_feature 13670. .13793
                /gene="WUGSC:H.RG083M05.1"
                /note="match to human 5' EST H41382 (NID:917434), bases
                143-266"
repeat_region 13794. .13877
                /rpt_family="ALU"
misc_feature 13878. .13906
                /gene="WUGSC:H.RG083M05.1"
                /note="match to human 5' EST H41382 (NID:g917434), bases
                30-58"
repeat_region 13907. .14104
                /rpt_family="ALU"
repeat_region complement(14110. .14137)
                /rpt_family="L1"
repeat_region complement(15618. .15907)
                /rpt_family="ALU"
repeat_region 17227. .17522
                /rpt_family="ALU"
misc_feature 18667. .19235
                /note="match to human fetal brain 5' EST D61494
                (NID:g970409), bases 1-255, and to human 3' EST R07476
                (NID:g759399)"
repeat_region 19550. .19670
                /rpt_family="ALU"
misc_feature 21507. .37303
                /note="similarity to various SS-RNA virus polypeptides;
                pseudogene, region of matches and close matches to
                multiple human ESTs, see R68740 (NID:g842257)"
exon 37316. .37489
                /note="Grali prediction, score = 80"
                /evidence=not experimental
repeat_region complement(38938. .39224)
                /rpt_family="ALU"
misc_feature 39225. .39707
                /note="match to multiple human ESTs, see N30113
                (NID:g1148633)"
repeat_region 39800. .40085
                /rpt_family="ALU"
repeat_region complement(40247. .40538)
                /rpt_family="ALU"
repeat_region complement(40632. .40924)
                /rpt_family="ALU"
repeat_region complement(42283. .42891)
                /rpt_family="ALU"
repeat_region complement(45474. .45613)
                /rpt_family="ALU"
misc_feature complement(45614. .45737)
                /note="match to human 3' EST H48898 (NID:g988738), bases
                129-333"
misc_feature complement(46107. .47026)
                /note="match to multiple human ESTs, see N81064
                (NID:g1243765), H48897 (NID:g988737), and M78831
                (NID:g273146)"

```

```

repeat_region complement(47027. .47318)
                /rpt_family="ALU"
misc_feature complement(47365. .47782)
                /note="match to multiple human ESTs, see W37495
                (NID:g1319089)"
misc_feature 47898. .48115
                /note="match to human 5' EST H62306 (NID:g1015138), bases
                93-368"
repeat_region complement(48116. .48405)
                /rpt_family="ALU"
misc_feature complement(48406. .48584)
                /note="match to human 3' EST N29952 (NID:g1148472), bases
                250-455, and 5' EST R12730 (NID:g765806)"
repeat_region complement(48787. .49405)
                /rpt_family="ALU"
misc_feature complement(49406. .49534)
                /note="match to human 3' EST R65794 (NID:g838432), bases
                309-440"
repeat_region complement(49638. .49672)
                /rpt_family="ALU"
misc_feature complement(49674. .49890)
                /note="match to human 3' EST N29952 (NID:g1148472) and 5'
                EST N29938 (NID:g1148458), sequences are from opposite
                ends of the same clone"
gene complement(49698. .51806)
                /gene="WUGSC:H.RG083M05.2"
                /complement(join(49698. .49888, 51575. .51806))
                /note="WUGSC:H.RG083M05.2"
                /note="coded for by human cDNAs W37389 (NID:g1319205),
                R65891 (NID:g838529), R65794 (NID:g838432) and R65794
                (NID:g838432)"
                /codon_start=1
                /protein_id="AAB46345.1"
                /cd_xref="GI:1669370"
                /translation="MFYRFQCGIIFCFPGVYVQGDVSVDEDDGKRYRQIRGF
                IDQYCEKSAALTMLPIPLSSPDQDPASVYIIGPEDLPFRKMEYLFVCHAPSEYRK
                SRSSPPVPTREPEKGIWTHVGPPTAIITIKESVANHL"
exon complement(51576. .51758)
                /gene="WUGSC:H.RG083M05.2"
                /note="Grali prediction, score = 86"
                /evidence=not experimental
repeat_region complement(52052. .52329)
                /rpt_family="L1"
misc_feature 55557. .55843

Query Match      85.0%; Score 6448; DB 9; Length 56093;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

```


QY 966 AAGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTCCATAATTCCTCAATTATGA 1025
DB 31313 AAGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTCCATAATTCCTCAATTATGA 31372
QY 1026 CCCCTCCCAAGACGTGGAGGAAAGAGAAATTCGGCCCAAGCCAGAGTGCATGTCTTTTXX 1085
DB 31373 CCCCT-CCAAAGCAGTGGAGGAAAGAAATTCGGCCCAAGCCAGAGTGCATGTCTTTTXX 31431
QY 1086 TCTCCAGACTTAAAGCAATATAAACAAGACTAGGTTAAATTCAGATTAATCTGATG 1145
DB 31432 TCTCCAGACTTAAAGCAATATAAACAAGACTAGGTTAAATTCAGATTAATCTGATG 31491
QY 1146 CTATATGTGTGTTTAAAGAGGTTAGAGCAATTCCTTGAATCTGACATGAGAGATATATA 1205
DB 31492 CTATATGTGTGTTTAAAGAGGTTAGAGCAATTCCTTGAATCTGACATGAGAGATATA 31550
QY 1206 TGTCACTGTAAATCAGACACTAACCCCAATGAGAGAGTGCACCACTAACCTGACGCT 1265
DB 31551 TGTCACTGTAAATCAGACACTAACCCCAATGAGAGAGTGCACCACTAACCTGACGCT 31610
QY 1266 GAGAGTTTGGCGATCTGTGTAATCTCAGTCAAGTCAATGATATGATGACAAACAGAG 1325
DB 31611 GAGAGTTTGGCGATCTGTGTAATCTCAGTCAAGTCAATG-ATAGATGACAAACAG-AGG 31667
QY 1326 AAAAGAAATGATTTCCCAAGCCAGACAGAGAGTCCAGTCTAACCCTCATTTGGGAG 1385
DB 31668 AAAAGAAATGATTTCCCAAGCCAGACAGAGAGTCCAGTCTAACCCTCATTTGGG--- 31724
QY 1386 ACAGAAATCAGTAACATGGGAGATTTGTCTGACACATTTGCTAACTTGTGTCTASAA 1445
DB 31725 ACAGAAATCAGAAACAT-AGAAATTTGTCTGACACATTTGCTAACTTGTGTCTASAA 31783
QY 1446 GGACTAAGGAAATCTASGAAAGAAATCTAYGAATTCAGATGATGTCACCACTAACACA 1505
DB 31784 GGACTAAGGAAATCTASGAAAG-AGCTCTATGAAATTCAGATGATGTCACCACTAACACA 31842
QY 1506 GGGGAGGAGGAAATCTCTACTGCTTTCTGAGAGACTAAAGGAGGACTTGAAGAGC 1565
DB 31843 -GGGAGGAGGAAATCTCTACTGCTTTCTGAGAGACTAAAGGAGGACTTGAAGAGC 31901
QY 1566 GTGCTCTCTGACCTGACTCTCTTCTGAAGGCCACTAATCTTAAAGCTAAGTTTATA 1625
DB 31902 GTGCTCTCTGACCTGACTCTCTTCTGAAGGCCACTAATCTTAAAGCTAAGTTTATA 31961
QY 1626 CTCACTGACGCTCAGACATTAG-AAAAAATCTCAAAAGTCTGCGGAGGCCGAGACAA 1684
DB 31962 CTCACTGACGCTCAGACATTAGAAAAAATCTCAAAAGTCTGCGGAGGCCGAGACAA 32021
QY 1685 ACTTGAAGAACCTTATGAACTTGGCAACTGCGTTTATTAATAGATCAGAGAGAC 1744
DB 32022 ACTTGAAGAACCTTATGAACTTGGCAACTGCGTTTATTAATAGATCAGAGAGAC 32081
QY 1745 AGGCGGAAACAGACAAACGGGATTTAAAAAAGGCCACGCTTTAGTCAATGACCTCAG 1804
DB 32082 AGGCGGAAACAGACAAACGGGATTTAAAAAAGGCCACGCTTTAGTCAATGACCTCAG 32141
QY 1805 CAAGTGAATTTGAGGCTCTGAGAAAGGAGAAAGTGGGCAATGATGCTAATAG 1864
DB 32142 CAAGTGAATTTGAGGCTCTGAGAAAGGAGAAAGTGGGCAATGATGCTAATAG 32201
QY 1865 GCTTTCCTTCAAGTGGCTCTCAAGGACACTTTAAAAAAGTTGTCAGATGAGAGTAA 1924
DB 32202 GCTTTCCTTCAAGTGGCTCTCAAGGACACTTTAAAAAAGTTGTCAGATGAGAGTAA 32261
QY 1925 CCGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGAGAGGCCACTGCCCCAG 1984
DB 32262 CCGCCCTTCTGTCATGCCCCCTTATTTCAAGGGAATCACTGAGAGGCCACTGCCCCAG 32321
QY 1985 GGAACAAGGCTTTTGAAGTCAAGGCACTAACCCAGATGATCCAGAGAGAGAGCTGAGG 2044
DB 32322 GGAACAAGGCTTTTGAAGTCAAGGCACTAACCCAGATGATCCAGAGAGAGAGCTGAGG 32381
QY 2045 TGCTGTGGGCAAGCCCATCCATGCCATCACTCTCAGAGCCCTGTGATGCTTGAAC 2104

DB 32382 TGCTGTGGGCAAGCCCATCCATGCCATCACTCTCAGAGCCCTGTGATGCTTGAAC 32441
QY 2105 ATTGAGGGCCAGGAGTGTCTCTGACACTGAGTGGGCTTCTTACTTACTCTTC 2164
DB 32442 ATTGAGGGCCAGG-AGGTTGTCTGACACTGAGTGGGCTTCTTACTTACTCTTC 32500
QY 2165 TGTCCCGAACACTGCTCTCCAGATCTGTCACTATTCGTGAGGGGGTCCCTTAAGACGGGCA 2224
DB 32501 TGTCCCGAACACTGCTCTCCAGATCTGTCACTA-TCGAGGGGGTCC-TCAGAGGGCA 32558
QY 2225 GTCACTAATCTTCTTCCAGCCACTAAGTATGAACTGGGAGCTTATCTTTTCA 2284
DB 32559 GTCACTAATATC-TTCTCCAGCCACTAAGTATG-ACGGGAGCTTATCTTTTCA 32616
QY 2285 ATGCTTTCTAATATGCTGAAAGCCCACTACCTGTTAGGAGAGCACTTACGAA 2344
DB 32617 ATGCTTTCTAATATGCTGAAAGCCCACTACCTGTTAGGAGAGCACTTACGAA 32676
QY 2345 AAGCAGGGCCATTATACACTGAAACATAGAGAGAAACACCCGTTGTGTGTCCTG 2404
DB 32677 AAGCAGGGCCATTATACACTGAAACATAGAGAGAAACACCCGTTGTGT-CCCTG 32735
QY 2405 CTGAGAAAGGATTAATCTGAACTCTGGGCAACAGAAAGCAATATGACAGGCCAA 2464
DB 32736 CTGAGAAAGGATTAATCTGAACTCTGGGCAACAGAAAGCAATATGACAG-CAAA 32794
QY 2465 GAATGCCGCTCTGTTCAAGTTAACTAAAGATTCACCTTCCCTACCAAGGCA 2524
DB 32795 GAATGCCGCTCTGTTCAAGTTAACTAAAGATTCACCTTCCCTACCAAGGCA 32854
QY 2525 GTACCCCTCAGACCCCAAGGCCCAACAGAAATTCMAAAGATTTGTTAAGACTTAAAGC 2584
DB 32855 GTACCCCTCAGACCCCAAGGCCCAACAGAAATTCMAAAGATTTGTTAAGACTTAAAGC 32914
QY 2585 CCAAGGCTTAAATTAACATGATTAATCTCTGACGTAATTCGATGAGATTGAGAGAG 2644
DB 32915 CCAAGGCTTAAATTAACATGATTAACCTGACGTAATTCG-ATTTTAAAGAG 32967
QY 2645 CACAGAAACCAAGTGAACAGTGAAGGAGTTAGTGAAGATTCAGAGTTATCAATGAGG 2704
DB 32968 TACAGAAACCAACAGAGAGTGA-AGTTAGTGAAGATTCAGAGTTATCAAT-AGGCG 33025
QY 2705 CGTTGCTCTTAAATCCAGCTGTAACCTTAACTGTGMYTTCCTCAATACAGA 2764
DB 33026 TGTTGCTCTTAATGACAGCTGTAACCTTAACTGTGMYTTCCTCAATACAGA 33085
QY 2765 GGAAGCAGAGTGTTAACATCTGAGACTTMAAGATGCTTCTGATCCCTGTACA 2824
DB 33086 GGAAGCAGAGTGTTAACATCTGAGACTTMAAGATGCTTCTGATCCCTGTACA 33145
QY 2825 TCTGACCTCAATCTTGTGCTTGAAGATCTTCAAAACCAACATCTCACTCAC 2884
DB 33146 TCTGACCTCAATCTTGTGCTTGAAGATCTTCAAAACCAACATCTCACTCAC 33205
QY 2885 CTGACCTTTTAAACCAAGGTTCAAGGATAGTCCCATATTTTGGCAGGCACTTAC 2944
DB 33206 CTGACCTTTTAAACCAAGGTTCAAGGATAGTCCCATATTTTGGCAGGCACTTAC 33265
QY 2945 CCAAGACTTGAAGCAATCTCATCTGAGACA-CTTCTCTCTGCTGAGTGAATTT 33323
DB 33266 CCAAGACTTGAAGCAATCTCATCTGAGACA-CTTCTCTCTGCTGAGTGAATTT 33383
QY 3005 ACTTTTTCGCGCTTCAAGAACTTGTGATCAACCAAGGCTCTTAAATTT 3064
DB 33324 ACTTTTTCGCGCTTCAAGAACTTGTGATCAACCAAGGCTCTTCAATTT 33383
QY 3065 CCTGCTACCTGTGCTACAGGTTTCCAAACGATCTCTGCTCACAGAGT 3124
DB 33384 CCTGCTACCTGTGCTACAGGTTTCCAAACGATCTCTGCTCACAGAGT 33443
QY 3125 TAAATCTTAAAGCTAATATTCAAAGGCAACAGGCTCACTGAGAGAAATATCA 3184

Db 33444 ----TACTTAGGGCTAAATTTATCCAAAGCACCAGGCCCTCAGTGAAGACACATCCA 33499
Qy 3185 GCCTATACTGGCTTATTCCTCATCYCAAAACCTTAAGCACTAAGRGRTTCTTGGCCT 3244
Db 33500 GCTTATACGGCTTATTCCTCATCCCAAAACCTTAAGCACTAAGGGGATTCCTTGGCGT 33559
Qy 3245 AAYAGGTTCTGCCGAAMATGATTCGCCAGTMTGGCRAAATAGCCAGTCATTMATATA 3304
Db 33560 AATAGGTTCTGCCGAAMATGATTCGCCAGTMTGGCRAAATAGCCAGTCATTMATATA 33618
Qy 3305 CASTAATTAAGAACTCAGAAAGCCATACCACTTAATTAAGTGAATGATGAGTGAAGTGA 3364
Db 33619 CATTAATTAAGAACTCAGAAAGCCATACCACTTAATTAAGTGAATGATGAGTGAAGTGA 33678
Qy 3365 AAGTGCTTTCCAGGCCCTTAAGAAAGCCTTAAGCCCAAGCCCAAGTCTTAAGTTC 3424
Db 33679 AAGTGCTTTCCAGGCCCTTAAGAAAGCCTTAAGCCCAAGCCCAAGTCTTAAGTTC 33724
Qy 3425 AACGGGCAAGACTTTTSTTATATRTACAGAAATAAAGRAAATGCTTGGAGTCC 3484
Db 33725 AACAGGCAAGACTTTTCTATATGTCACAG-AAAAAAGAGAAATGCTTGGAGTCC 33783
Qy 3485 TTAACAGRTCCAGAGGAYAGGCTTGCAACCTGTGCGRACTGASTPAGAAATGATG 3544
Db 33784 TTAACAGRTCCAGAGGATGAGCTTGCAACCTGTGCGRACTGASTPAGAAATGATG 33843
Qy 3545 TAGTGCAAAAGGGTTGRCYTCATTTTAYGGGATGATGAGTGAAGTGAAGTGAATAT 3604
Db 33844 TAGTGCAAAAGGGTTGRCYTCATTTTAYGGGATGATGAGTGAAGTGAAGTGAATAT 33903
Qy 3605 CTGAAGCAGTTAAATATATACAGGAGAGATCTTACTGTGTGACATCTCATGAKTGA 3664
Db 33904 CTGAACACAGTTAAATATATACAGGAGAGATCTTACTGTGTGACATCTCATGAGTGA 33963
Qy 3665 AYRGCTACTCACCTGCTTAAGAGAGCTTGTGGCTGTCAGACCAAGTCTTAATATTC 3724
Db 33964 ATGGCTACTCACCTGCTTAAGAGAGCTTGTGGCTGTCAGACCAAGTCTTAATATTC 34023
Qy 3725 AGGCTCTATTAAGTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTTTAACCCAG 3784
Db 34024 AGGCTCTATTAAGTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTTTAACCCAG 34083
Qy 3785 YNCATTTTCTTCCAGACAAATGAAGAAAGATPAAATCTGTGCAACTTTAAATTTCTC 3844
Db 34084 CCAATTTTCTTCCAGACAAATGAAGAAAGATPAAATCTGTGCAACTTTAAATTTCTC 34143
Qy 3845 AAACCTATGCACTCGAGGGGACCTTGTAGAGTTCCTTGAAGTCTGACCTGACCTTGAAC 3904
Db 34144 AAACCTATGCACTCGAGGGGACCTTGTAGAGTTCCTTGAAGTCTGACCTGACCTTGAAC 34202
Qy 3905 TTGTATACTGATGAGAGTTCCTTTGTAGAAAGAGCTTGAAGAGTGGGTATGAGTG 3964
Db 34203 TTGTATACTGATGAGAGTTCCTTTGTAGAAAGAGCTTGAAGAGTGGGTATGAGTG 34262
Qy 3965 GTCAGTGATATGAGAAATTTGAAGTATCCCTCACTCGAGAACTAAGTCTTAACTR 4024
Db 34263 GTCAGTGATATGAGAAATTTGAAGTATCCCTCACTCGAGAACTAAGTCTTAACTR 34322
Qy 4025 GCGAATTAATAGCCYCAATKGGGCACTGAATTTAGAGAGAGAAAGGGVAAATATA 4084
Db 34323 GCGAATTAATAGCCYCAATKGGGCACTGAATTTAGAGAGAGAGAAAGGGVAAATATA 34382
Qy 4085 TATACAGACTCTAATATGCTAATGCTAGTCTTCATGCGCAATGAGCAATATGAGAA 4144
Db 34383 TATACAGACTCTAATATGCTAATGCTAGTCTTCATGCGCAATGAGCAATATGAGAA 34442
Qy 4145 AGGGAATTCCTAATCTTGAAGAGAACCTATCAACATCGAGAAAGCCATTAGAGATTA 4204
Db 34443 AGGGAATTCCTAATCTTGAAGAGAACCTATCAACATCGAGAAAGCCATTAGAGATTA 34502
Qy 4205 TTATGTGCTGACAGAAACCTAAGAGTGTGAGTCTTACATCTGCGGGGTCACTACAGAA 4264
Db 34503 TTATGTGCTGACAGAAACCTAAGAGTGTGAGTCTTACATCTGCGGGGTCACTACAGAA 34562

Qy 4265 GGAAGRAAAGGAAATATAGAAATTTGCCAAGCAKATTTGAAGCAAAAGAGTGC 4324
Db 34563 GGAAGRAAAGGAAATATAGAAATTTGCCAAGCAKATTTGAAGCAAAAGAGTGC 34622
Qy 4325 AGGAGAGACCTCCATTAAGAAATGCTTATTAACCTTCCCTAGTATAGGGTAAATCCCTTC 4384
Db 34623 AGGAGAGACCTCCATTAAGAAATGCTTATTAAGAAATCCCTAGTATAGGGTAAATCCCTTC 34682
Qy 4385 CGGAAACCAAGCCCAAGTACTCAGAGAGAAACAGAAATGGGAAACCTCAGAG- CAG 4443
Db 34683 CGGAAACCAAGCCCAAGTACTCAGAGAGAAACAGAAATGGGAAACCTCAGAGAG- CAG 34742
Qy 4444 TTTTCTCCCTGGGAGCGGTTAGCCATGAGAGAGGAAATATCTTTGCTGCACTAT 4503
Db 34743 TTTTCTCCCTGGGAGCGGTTAGCCATGAGAGAGGAAATATCTTTGCTGCACTAT 34802
Qy 4504 CCAATGGAATTTACTTAACCTTTCATCAACCTTTCACTTAAGCATGATGACCA 4563
Db 34803 CCAATGGAATTTACTTAACCTTTCATCAACCTTTCACTTAAGCATGATGACCA 34862
Qy 4564 TCATATGCTCAATCATTTATTAATGACCAAGGCTTTCAAAATCATCAAGCAATAT 4623
Db 34863 TCATATGCTCAATCATTTATTAATGACCAAGGCTTTCAAAATCATCAAGCAATAT 34922
Qy 4624 CAGGCTGTGAATGTCGACABAAATATCCCTGCTATCGCAAGCTCTTCAG 4683
Db 34923 CAGGCTGTGAATGTCGACABAAATATCCCTGCTATCGCAAGCTCTTCAG 34982
Qy 4684 ARAACAAABAACAGGCACTTAACCTGTRABABAACCTGCAATTTTACCAACAGCC 4743
Db 34983 ARAACAAABAACAGGCACTTAACCTGTRABABAACCTGCAATTTTACCAACAGCC 35042
Qy 4744 AAACCTCAGGATTTCAATATCTACTGTGCTGATATATCTTTACAGGTTGGGCAAG 4803
Db 35043 AAACCTCAGGATTTCAATATCTACTGTGCTGATATATCTTTACAGGTTGGGCAAG 35102
Qy 4804 GCCTTCCCTGTAGAGCAGAAAGGCCAAGAGTAAAGGCACTAGTTCAATGAATA 4863
Db 35103 GCCTTCCCTGTAGAGCAGAAAGGCCAAGAGTAAAGGCACTAGTTCAATGAATA 35162
Qy 4864 ATTCCCAATTCGACTTCCCGAGGCTTAACAGATGACAAATAGCCCTGCTTTCAGGCC 4923
Db 35163 ATTCCCAATTCGACTTCCCGAGGCTTAACAGATGACAAATAGCCCTGCTTTCAGGCC 35222
Qy 4924 ACAATTAACCAAGGATATCCAGGCTTATGATATGATATCACTTAACATGCGGCTGA 4983
Db 35223 ACAATTAACCAAGGATATCCAGGCTTATGATATGATATCACTTAACATGCGGCTGA 35282
Qy 4984 AGGCCCAAGTCTCAGGAAAGGTCGAGAAATGAAGAAATCAAAAGGACATCTTAA 5043
Db 35283 AGGCCCAAGTCTCAGGAAAGGTCGAGAAATGAAGAAATCAAAAGGACATCTTAA 35342
Qy 5044 AAGCAAAACCAAGGAAACCACTCAATGCTGCTGTGCTATAGCTTAAAGAA 5103
Db 35343 AAGCAAAACCAAGGAAACCACTCAATGCTGCTGTGCTATAGCTTAAAGAA 35402
Qy 5104 ATCTGCACTTTCCCAAAAGCAGACTTAAGCCCATCAATATGCTGTATGAGAGGCC 5163
Db 35403 ATCTGCACTTTCCCAAAAGCAGACTTAAGCCCATCAATATGCTGTATGAGAGGCC 35462
Qy 5164 TTCAATTAACCAATGACTTGTGCTTGAACCAAGCAAGCACTTATGTTGAGACATCACT 5223
Db 35463 TTCAATTAACCAATGACTTGTGCTTGAACCAAGCAAGCACTTATGTTGAGACATCACT 35522
Qy 5224 CTTAGCAATATCAACAAAGTCTTAAATCAATTAACAGAACTATCCCTGAGAAAGG 5283
Db 35523 CTTAGCAATATCAACAAAGTCTTAAATCAATTAACAGAACTATCCCTGAGAAAGG 35582
Qy 5284 GAAAGAACTATTCACCCCTGAGCAATGATATTAAGTCAAGTCCCTTCCTATATTC 5343
Db 35583 GAAAGAACTATTCACCCCTGAGCAATGATATTAAGTCAAGTCCCTTCCTATATTC 35642

QY	53444	CATCCCTGATACATCTCGGGAGGAGCCCTACCCAGCATCTTTATATACCCCAATCGCCG	5403
Db	35643	CATCCCTGATACATCTCGGGAGGAGCCCTACCCAGCATCTTTATATACCCCAATCGCCG	35706
QY	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATATCATCACTTGAGTGAATAATCCGTGATACGC	5463
Db	35703	TTAAAGTGGCTGGAGTGGAGTCTTGGATATCATCACTTGAGTGAATAATCCGTGATACGC	35766
QY	5464	CAAGGAACCTGAAAAATCCAGAGACAACGTTAGCTATTCTGTGAACCTCTAGAGATT	5523
Db	35763	CAAGGAACCTGAAAAATCCAGAGACAACGTTAGCTATTCTGTGAACCTCTAGAGATT	35822
QY	5524	TGGCGTCTCTTCAACAACAACAGAGAAAGTAACTTAAATTCATAATCCCCANG	5583
Db	35823	TGGCGTCTCTTCAACAACAACAGAGAAAGTAACTTAAATTCATAATCCCCANG	35881
QY	5584	GSCTCCCTTATCATATTTTCTCTKTAISTGTSSTTVAACCTSTTTCACCTCACTGCA	5643
Db	35882	GCCCTCCCTTATCATATTTTCTCTTACTCTTCTTTAACCTCTTTCACCTCACTGCA	35944
QY	5644	CCCCCTCATGCGCTGATACCAAGTAGCTCCCTTACCMAGATTCTATGAGAATG	5703
Db	35944	CCCCCTCATGCGCTGATACCAAGTAGCTCCCTTACCMAGATTCTATGAGAATG	36001
QY	5704	CAGCCTCCCCGAAAAATTTGATGCCCCCATCGTATAGAGTCTTTSYAAGGAAACCCCAAC	5763
Db	36002	CAGCCTCCCCGAAAAATTTGATGCCCCCATCGTATAGAGTCTTTCYAGGGAAACCCCAAC	36061
QY	5764	TTCACTGCGCCACACCCATATGCCCGCAACGCTATCACTGCGCACTCTTSGCATGCAT	5823
Db	36062	TTCACTGCGCCACACCCATATGCCCGCAACGCTATCACTGCGCACTCTTSGCATGCAT	36121
QY	5824	GCAAACTACTTATTGGACAGAAAAATGATTATCCTAGTTGTCCGTGAGACCTTGA	5883
Db	36122	GCAAACTACTTATTGGACAGAAAAATGATTATCCTAGTTGTCCGTGAGACCTTGA	36181
QY	5884	GTCATCTGCTGTGGACTTACTTCAACCCAACTGTATGTCTGATGGGGGTGAATTCAA	5943
Db	36182	GTCATCTGCTGTGGACTTACTTCAACCCAACTGTATGTCTGATGGGGGTGAATTCAA	36241
QY	5944	GATCAGGCAAGAAAAACATGTAAAAAGAGTATCTCCCACTCACCGSGGGATCATGGC	6003
Db	36242	GATCAGGCAAGAAAAACATGTAAAAAGAGTATCTCCCACTCACCGSGGGATCATGGC	36301
QY	6004	ACCTTACGCCCTTCAAAAGACTGATCTCTCAAACTAATGAAACCCCTCGTACCAT	6063
Db	36302	ACCTTACGCCCTTCAAAAGACTGATCTCTCAAACTAATGAAACCCCTCGTACCAT	36361
QY	6064	ACTGCGCTGTAAAGCTATTTAATACACCCCTCACTGSGGCTCATGAGTCTCGGCCAA	6123
Db	36362	ACTGCGCTGTAAAGCTATTTAATACACCCCTCACTGSGGCTCATGAGTCTCGGCCAA	36421
QY	6124	AAACCTACTACTGTGGATATGCTCCCTCGTGAACCTTCARGCATATGTTTCAATCCCT	6183
Db	36422	AAACCTACTACTGTGGATATGCTCCCTCGTGAACCTTCARGCATATGTTTCAATCCCT	36481
QY	6184	GTAACCTGAACATGGAACAACCTTCAGACAGAAATTAACAACACTTCCGTTTATGAGGA	6243
Db	36482	GTAACCTGAACATGGAACAACCTTCAGACAGAAATTAACAACAACACTTCCGTTTATGAGGA	36541
QY	6244	CCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTTAGC	6303
Db	36542	CCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTTAGC	36601
QY	6304	AATATCTCATACACAACCACTCCCAATGCAATCAGTGGGTAATCTCTCCCAACAATAA	6363
Db	36602	AATATCTCATACACAACCACTCCCAATGCAATCAGTGGGTAATCTCTCCCAACAATAA	36661
QY	6364	GTCGACCAACCTCAGGAATATTTTGTGTGTGTGTCTCAACCTATCGTGTGTAAT	6423
Db	36662	GTCGACCAACCTCAGGAATATTTTGTGTGTGTGTCTCAACCTATCGTGTGTAAT	36721
QY	6424	GGCCTTCAGATCTATGTGCTTCTCTCATTTTACTGCCCCCAATGRCATCTACCT	6483

Db	36722	GGCTTCAGATCTAATGCTTCCTCAATCTTAGTGGCCCCCTTAGACATCTACACT	36781
Qy	6484	GAACAAGATTATACAGTTATGTCATATCTAAGCCCCGCAACAAAGAGTACCATTCTT	6543
Db	36782	GAACAAGATTATATACATTATGTCATATCTAAGCCCCGCAACAAAGAGTACCATTCTT	3684
Qy	6544	CCCTTTGTTATAGGACGAGAGTGTAGTGCATAGAGTCTGGCATTTGGCGGTATCA	6603
Db	36842	CCCTTTGTTATAGGACGAGAGTGTAGTGCATAGAGTCTGGCATTTGGCGGTATCA	36901
Qy	6604	ACCTCTACTAGTCTCTCTCAAACTATCTCAAGAACTAAATGGGGACATAGAACGGGTC	6663
Db	36902	ACCTCTACTAGTCTCTCTCAAACTATCTCAAGAACTAAATGGGGACATAGAACGGGTC	36966
Qy	6664	GCCGACTCCCTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGCAGTACTCTTCA	6723
Db	36962	GCCGACTCCCTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGCAGTACTCTTCA	37022
Qy	6724	AATGAAAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAACTGTATTATTTAGG	6783
Db	37022	AATGAAAGCTTTAGACTYGTCTAACCGCTGAAAGAGGGGAACTGTATTATTTAGG	37081
Qy	6784	GAAGAATGCTGTTATATGTTATTCATCCGGAATGTCACTGAGAGAACTTTAAGAAAT	6843
Db	37082	GAAGAATGCTGTTATATGTTATTCATCCGGAATGTCACTGAGAGAACTTTAAGAAAT	37141
Qy	6844	CSAGATCGAATATACAGCTATAGAGAGAGCTTCGAAACACTGGAACCTTGCGGCTCTC	6903
Db	37142	CGAATTCGAATACAGCTATAGAGAGAGCTTCGAAACACTGGAACCTTGCGGCTCTC	37201
Qy	6904	AGCCRATGATGCTCGTGGATCTCCCTCTTAGACCTCTAGACCTATATATTGCTA	6963
Db	37202	AGCCRATGATGCTCGTGGATCTCCCTCTTAGACCTCTAGACCTATATATTGCTA	37261
Qy	6964	CTCTCTTTTGGACCTGTATCTTTTACCTCTGTTTAACTTTGTCTCTTCCAGATCGAA	7023
Db	37262	CTCTCTTTTGGACCTGTATCTTTTACCTCTGTTTAACTTTGTCTCTTCCAGATCGAA	37321
Qy	7024	GCTGTAAACATCAAAATGAGAGCCCAAGATGCACTGCAAGCTTAACGCGACAGCC	7083
Db	37322	GCTGTAAACATCAAAATGAGAGCCCAAGATGCACTGCAAGCTTAACGCGACAGCC	37381
Qy	7084	CTGACCGGCGCTGYTAGCCAGATCTGATGTTTATGATCATCAAGAGCACCCCTCTGAG	7143
Db	37382	CTGACCGGCGCTGYTAGCCAGATCTGATGTTTATGATCATCAAGAGCACCCCTCTGAG	37441
Qy	7144	GAATCTCAGCTGACAACTCTTACTACGCCCAATTGACGAGAAAGCAGTTAGAGCGT	7203
Db	37442	GAATCTCAGCTGACAACTCTTACTACGCCCAATTGACGAGAAAGCAGTTAGAGCGT	37501
Qy	7204	SGTGGGCAACCTCCCAACAGACTTAGGTTTCCGTGTGAGATGGGGGACGTAGAGAC	7263
Db	37502	SGTGGGCAACCTCCCAACAGACTTAGGTTTCCGTGTGAGATGGGGGACGTAGAGAC	37561
Qy	7264	AGGACTAGCTGAATTCCTAGGCTGATTAAGAACTCCYTAAAGCTAGTGGGAAGTGACC	7323
Db	37562	AGGACTAGCTGAATTCCTAGGCTGATTAAGAACTCCYTAAAGCTAGTGGGAAGTGACC	37621
Qy	7324	ACATCCACTTTTAAACGCGGGCTTGCACTTAATGTCACACTGACCAATCAGAGAGCTC	7383
Db	37622	ACATCCACTTTTAAACGCGGGCTTGCACTTAATGTCACACTGACCAATCAGAGAGCTC	37681
Qy	7384	ACTTAAATGCTAATTTAGGCAAAAGACGAGATTAAGAAATTAACATCATYATATGCMTG	7443
Db	37682	ACTTAAATGCTAATTTAGGCAAAAGACGAGATTAAGAAATTAACATCATYATATGCMTG	37741
Qy	7444	AGAGCAGACGAGAGGACATATGATGGGATATTAACCCCAAGTTCAGAGCGGCAACGG	7503
Db	37742	AGAGCAGACGAGAGGACATATGATGGGATATTAACCCCAAGTTCAGAGCGGCAACGG	37801
Qy	7504	CAACCCCTTTGGGTCCCTCCCTTTGATGGAAGCTCTGTTTCAATGCTATTTCACTCT	7563

Db 37802 CAACCCCTTGGGCTCCCTTCCTTGATGAGAGCTCTGTTCATGCTATTCTACTCT 37861

Qy 7564 ATTAATCTTGACCTGCR 7582

Db 37862 ATTAATCTTGACCTGCA 37880

RESULT 9
AC007566/c 149194 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.
DEFINITION AC007566
AC007566.2 GI:11181861
VERSION AC007566.2 GI:11181861
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 149194)
Suleston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
9847074
2 (bases 1 to 149194)
Du, Z.
The sequence of Homo sapiens BAC clone CTB-10G5
Unpublished (2001)
3 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (15-MAY-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (16-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (03-JUN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (06-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
8 (bases 1 to 149194)
Waterston, R.H.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 16, 2000 this sequence version replaced gi:14835815.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
Summary Statistics
Center project name: H_RG010G05

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send
<mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library
CTB-978SK-B. The library contains cloned DNA from the male
fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad.
Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8
(1996). This clone is available from Research Genetics, Inc.
(<http://www.reegen.com>).

VECTOR: pBelOBAC11

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-91H5, 200 base pair
overlap. Actual start of this clone is at base position 195 of
CTB-10G5; actual end is at base position 150532 of CTB-10G5.

The clone CTB-10G5 contains the entire sequence of CTB-83M5.

FEATURES

source

Location/Qualifiers

1..149194
/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q21-7q22"

/clone="CTB-10G5"

/clone_1ib="CTB-978SK-B"

1..1634

/rpt_family="L1"

2248..2388

/note="match to EST BG752883 (NID:g14063536)"

2248..2387

/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"

2248..2387

/note="match to EST AW579261 (NID:g7254310)"

2248..2387

/note="match to EST BG766882 (NID:g1407535)"

2248..2387

/note="match to EST B1160365 (NID:g14620366)"

2248..2374

/note="match to EST B1160365 (NID:g14620366)"

2248..2287

/note="similar to Homo sapiens EST BF758865
(NID:g12106765)"

2248..2287

/note="match to EST AV686676 (NID:g10288539)"

2253..2387

/note="match to EST AU123510 (NID:g10948226)"

2344..2387

/note="match to EST AV686676 (NID:g10288539)"

2696..3066

/rpt_family="MaLR"

3108..3392

/rpt_family="Alu"

3540..3628

misc_feature

repeat_region

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature	/note="match to EST AA425526 (NID:g2106267) zw48d03.r1"
	3540. .3628
	/note="similar to Mus musculus EST BB253526 (NID:g8946272) "
misc_feature	3542. .3632
	/note="match to EST AV686676 (NID:g10208939) "
misc_feature	3542. .3628
	/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
misc_feature	3542. .3628
	/note="match to EST AU123510 (NID:g10948226) "
misc_feature	3542. .3628
	/note="match to EST BG766882 (NID:g14077535) "
misc_feature	3542. .3628
	/note="match to EST B1160365 (NID:g14620366) "
misc_feature	3542. .3609
	/note="match to EST AW579261 (NID:g7254310) "
misc_feature	3562. .3628
	/note="match to EST BE272564 (NID:g9146913) "
misc_feature	3598. .3628
	/note="match to EST BE299708 (NID:g9183456) "
misc_feature	3714. .3785
	/note="match to EST BG260659 (NID:g12770475) "
misc_feature	3717. .3785
	/note="similar to Mus musculus EST BE94936 (NID:g10678674) "
misc_feature	3719. .3785
	/note="match to EST AA149693 (NID:g1720635) zn99d06.r1"
misc_feature	3719. .3785
	/note="match to EST AA425526 (NID:g2106267) zw48d03.r1"
misc_feature	3719. .3785
	/note="match to EST AU123510 (NID:g10948226) "
misc_feature	3719. .3785
	/note="match to EST BE272564 (NID:g9146913) "
misc_feature	3719. .3785
	/note="match to EST BE299708 (NID:g9183456) "
misc_feature	3719. .3785
	/note="match to EST BG766882 (NID:g14077535) "
misc_feature	3719. .3785
	/note="match to EST B1160365 (NID:g14620366) "
misc_feature	3719. .3785
	/note="similar to Mus musculus EST BB253526 (NID:g8946272) "
misc_feature	3860. .3862
	/note="match to EST AV686676 (NID:g10288539) "
repeat_region	4730. .4827
	/fpL family="L2"
repeat_region	4834. .4925
	/fpL family="L1"
misc_feature	5715. .5860
	/note="match to EST AA425526 (NID:g2106267) zw48d03.r1"
misc_feature	5715. .5860
	/note="match to EST AU123510 (NID:g10948226) "
misc_feature	5715. .5860
	/note="match to EST BE272564 (NID:g9146913) "
misc_feature	5715. .5860
	/note="match to EST BE299708 (NID:g9183456) "
misc_feature	5715. .5860
	/note="match to EST BG260659 (NID:g12770475) "
misc_feature	5715. .5860
	/note="match to EST BG766882 (NID:g14077535) "
misc_feature	5715. .5860
	/note="match to EST B1160365 (NID:g14620366) "
misc_feature	5715. .5860

Query Match	Similarity	85.0%	Score 6448	DB 9	Length 149194
Best Local	Similarity 95.7%	Pres. No. 0			
Matches 6681	Conservative 195	Mismatches 50	Indels 53	Gaps 27	
Oy	606	GATGGGAAACCTTCCCGGCAAGACAAAAGCCGCTTAAGACGTAATTGGAATAATTGGGA	665		
Db	91051	GATGGGAAACCTTCCCGGCAAGACAAAAGCCGCTTAAGACGTAATTGGAATAATTGGGA	90925		
Oy	666	MCAATTGACCCCTCAGACACTAAGAAAGAAAGCACTTAATTCTTCGACAGTCGCGCTT	725		

Db	90991	CCAAATTGACCTCTAGACACTTAAGAAAACGACCTTATATTTCTTCGACGTGCGGCTG	90993
Qy	726	GCACCTCCGAGGGAGATATAAATTATAACCACTCTTACAGCTAGACCTTTTGTGAGAA	785
Db	90991	GCACTCCTGAGGGAGATATAAATTATAACCACTCTTACAGCTAGACCTTTTGTGAGAA	9087
Qy	786	AAGCCAAATGAGTGAAGTGCCATTAAGTACAACTTTCTTTTCATTAAGACAACTCAC	845
Db	90871	AAGCCAAATGAGTGAAGTGCCATTAAGTACAACTTTCTTTTCATTAAGACAACTCAC	9081
Qy	846	AATTATGTAAAAATGTGATTTTATGCCCTTACAGAAAGCTTTCAGAGCTACCTCCCTATC	905
Db	90811	AATTATGTAAAAATGTGATTTTATGCCCTTACAGAAAGCTTTCAGAGCTACCTCCCTATC	9075
Qy	906	CCAGCATCCCGGACCTCCCTCCCGCAATTAAGAACCCCTCTCAACCCAAATGATGCCAA	965
Db	90751	CCAGCATCCCGGACCTCCCTCCCGCAATTAAGAACCCCTCTCAACCCAAATGATGCCAA	9069
Qy	966	AAGAGATTAAGCAAAAAGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATTATGA	1025
Db	90691	AAGAGATTAAGCAAAAAGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAATTATGA	90631
Qy	1026	CCCTTCCCAAGCAGTGGGAGGAAGAAATTCGGCCGACGAGAGTGATGTCGCTTTT	1085
Db	90631	CCCTTCCCAAGCAGTGGGAGGAAGAAATTCGGCCGACGAGAGTGATGTCGCTTTT	9057
Qy	1086	TCTCCCAAGCTTAAGCAATTAACCAAGACTTAAGTAAATTTCTAGATTAATCTGATG	1145
Db	90572	TCTCCCAAGCTTAAGCAATTAACCAAGACTTAAGTAAATTTCTAGATTAATCTGATG	9051
Qy	1146	CTATATTTGTGTATTTCAGAGGTTAAGCAATTTCTTTGATCTGACATGAGAGATATATA	1205
Db	90512	CTATATTTGTGTATTTCAGAGGTTAAGCAATTTCTTTGATCTGACATGAGAGATATATA	9045
Qy	1206	TGTCACTGTCTAATCAGACACTAACCCCAATAGAGAAAGTGCCACATACTGACGCT	1265
Db	90453	TGTCACTGTCTAATCAGACACTAACCCCAATAGAGAAAGTGCCACATACTGACGCT	9039
Qy	1266	GAGGTTTGGGAGTCTGTGATCTCAGTCAAGTCAATGATATNGATGCAACAGAAAG	1325
Db	90393	GAGGTTTGGGAGTCTGTGATCTCAGTCAAGTCAATGATATNGATGCAACAGAAAG	9033
Qy	1326	AAAGANAATGATTTCCCAAGGCGACAGCAGTCTCCCACTTAAACCTCATTTGGGAGC	1385
Db	90336	AAAGANAATGATTTCCCAAGGCGACAGCAGTCTCCCACTTAAACCTCATTTGGGAGC	9028
Qy	1386	ACAGAAATCAGTAAACATGGAGATTGGTCTGACAGACTTTGCTAACTTGTGTCTASAA	1445
Db	90279	ACAGAGATCAGAAACATGGAGATTGGTCTGACAGACTTTGCTAACTTGTGTCTASAA	9022
Qy	1446	GGACTAAGGAAATCTAAGAAATCTAAGAAATCTCAATGATGCCAATAACACA	1505
Db	90220	GGACTAAGGAAATCTAAGAAATCTAAGAAATCTCAATGATGCCAATAACACA	9016
Qy	1506	GGGAGAGGAAAGAAATCTTACTGCTTTCTGAGAGACTTAAGAGAGCAATTGAGAGC	1565
Db	90161	GGGAGAGGAAAGAAATCTTACTGCTTTCTGAGAGACTTAAGAGAGCAATTGAGAGC	9010
Qy	1566	GTGCTCTCTGTGACCTGACTTTCTTGAAGGCCAATACTTAAAGCGTAAATTTATCA	1625
Db	90102	GTGCTCTCTGTGACCTGACTTTCTTGAAGGCCAATACTTAAAGCGTAAATTTATCA	9004
Qy	1626	CTCAGTCAAGTGCAGACATTAAG-AAAAAATTCAAAAGTGTGCGTAGGCCCCGAGCAAA	1684
Db	90042	CTCAGTCAAGTGCAGACATTAAG-AAAAAATTCAAAAGTGTGCGTAGGCCCCGAGCAAA	8998
Qy	1685	ACTTAAGAAACCTTATTTGAACCTTGCCAACTTCGGTTTTTTTATTAATGAGATCAGAGAGC	1744
Db	89982	ACTTAAGAAACCTTATTTGAACCTTGCCAACTTCGGTTTTTTTATTAATGAGATCAGAGAGC	89922
Qy	1745	AGCGGAAACGAGCAAAAGGATTAATAAAGGCCACCGCTTATGATGACCTTCAGG	1804

Db	89292	AGGGGAAACAAGACAACAGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG	89863
Oy	1805	CAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGCCAAATTGAAATGCTTAATAGG	1864
Db	89862	CAAGTGGACTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGCCAAATTGAAATGCTTAATAGG	89803
Oy	1865	GCTTGCTTCCAGTGCGGGTCTACAGAGACCTTTAAAAAATTGTCCTCAAGTAAAGTAAG	1924
Db	89802	GCTTGCTTCCAGTGCGGGTCTACAGAGACCTTTAAAAAATTGTCCTCAAGTAAAGTAAG	89743
Oy	1925	CCGCCCCCTTGCTCATGCCCCCTTATTTTCAAGGAATCATGTGAAGGCCCATGCCCCAGG	1984
Db	89742	CCGCCCCCTTGCTCATGCCCCCTTATTTTCAAGGAATCATGTGAAGGCCCATGCCCCAGG	89683
Oy	1985	GGACAAGAGCTTTTGGAGTCAAGACCACTAACAGATGATCCAGAGACAGAGACTGAGGG	2044
Db	89682	GGACAAGAGCTTTTGGAGTCAAGACCACTAACAGATGATCCAGAGACAGAGACTGAGGG	89623
Oy	2045	TGCTGTGGGGCAAGGCGCATCCCATGCGCATCACCCCTCACAGAGCCTGGTATGCTTGACC	2104
Db	89622	TGCTGTGGGGCAAGGCGCATCCCATGCGCATCACCCCTCACAGAGCCTGGTATGCTTGACC	89563
Oy	2105	ATTGAGGGCCAGGAAAGTTGTCTCTCTGGAACATGTGTGGGCTTTTCTTAAGTCACTCTTC	2164
Db	89562	ATTGAGGGCCAGG-AGGTTGTCTCTCTGGAACATGTGTGGGCTTTTCTTAAGTCACTCTTC	89504
Oy	2165	TGTCCCGGACAACATGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCA	2224
Db	89503	TGTCCCGGACAACATGTCCTCCAGATCTGTCACTA-TGTGAGGGGGTCC-TAAGACGGGCA	89446
Oy	2225	GTCACTAGATACTTTTCTCCAGCCACTAGTAAGTAAGACTGGGAGCTTTATTTCTTTTCC	2284
Db	89445	GTCACTAGATATC-TTCTCCAGCCACTAGTAAGTAAG-ActGGGAGCTTTATTTCTTTTCC	89388
Oy	2285	ATGCTTTTCTAATTATGTTGAAGGCCCACTACCTTGTTAGGAGAGACATCTCTGACAA	2344
Db	89387	ATGCTTTTCTAATTATGTTGAAGGCCCACTACCTTGTTAGGAGAGACATCTCTGACAA	89328
Oy	2345	AAGCAGGGGCAATTATACACCTGAAACATAGGAGAGGAAACCCGTTTGTTGCCCTTG	2404
Db	89327	AAGCAGGGGCAATTATACACCTGAAACATAGGAGAGGAAACCCGTTTGTTGT-CCCTTG	89269
Oy	2405	CTTGAGAGAGAAATTAATCTTGAAGCTGTGGGCAACAAGACAAATATGAGAGAGGCAAA	2464
Db	89268	CTTGAGAGAGAAATTAATCTTGAAGCTGTGGGCAACAAGACAAATATGAGAGAG-CAAA	89210
Oy	2465	GAATGACCGGTCTGTTCAGTTAACTTAAGAGATTCACATCTCTTCCCTCAACMAAGGCA	2524
Db	89209	GAATGCCGCTCTGTTCAGTTAACTTAAGAGATTCACATCTCTTCCCTCAACMAAGGCA	89150
Oy	2525	GTACCCCTCAAGACCCCAAGGCCCAAGAGATTCCAAAAGATTGTTAAGACTTTAAAGC	2584
Db	89149	GTACCCCTCAAGACCCCAAGGCCCAAGAGACTCCAAAAGATTGTTAAGACTTTAAAGC	89090
Oy	2585	CCAAAGGCTTAAGTAAAAACATGCATTAATCTCCCTGACATTAATCCGTAGTGATGAGAGG	2644
Db	89089	CCAAAGGCTTAAGTAAAAACATGCATTAATCTCCCTGACATTAATCCGTAGTGATGAGAGG	89033
Oy	2645	CACAGAAAACCAAGTGGACAGTGAAGGAGTTAGTGCAAGATCTCAGAGATTAACTGAGGC	2704
Db	89036	TACAGAAAACCAAGTGGACAGTGAAGGAGTTAGTGCAAGATCTCAGAGATTAACTAAT-GAGGC	88979
Oy	2705	CGTTGTCTTTTATACCCAGCTGTATCCTAAGCCCTTAATCTGTGTTCCTCCAAATACAGA	2764
Db	88978	TGTTTGTCTCTATATGCAAGCTGTATCCTAAGCCCTTAATCTGTGTTCCTCCAAATACAGA	88918
Oy	2765	GGAAACAGATGTTTACATCTCTGACCTTMAAGATGCTTTCTTCTGATCCCTGTACA	2824
Db	88918	GGAAACAGATGTTTACATCTCTGACCTTCAAGATGCTTTCTTCTGATCCCTGTACA	88855
Oy	2825	TCCTGACTCTCAATTTCTTGTTGCTTTGAAGATCTTCAAAACCAACATCTCACTCAC	2884
Db	88858	TCCTGACTCTCAATTTCTTGTTGCTTTGAAGATCTTCAAAACCAACATCTCACTCAC	88795

QY	2885	CTGACATCTTTTATCCCAAGGGGTCAAGGATAGYCCCATCTATTTGGACAGGACTTATGC	2944
Db	88798	CTGACATATTTTATCCCAAGGGGTCAAGGATAGTCCCATCTATTTGGACAGGACTTATGC	88739
QY	2945	CCAAGACTTGAGYCAATTTTCATCTACCTGACACTCTTGCTCTTCGTGATGATGATTT	3004
Db	88738	CCAAGACTTGAGCAATCTCTCATCTGAGCA--CTTGCTCTTCGTGATGATGATTT	88681
QY	3005	ACTTTTGGCCGCTTTCAGAAACCTTGCCATCAGCAACCAAGGCTCTTCAATTT	3064
Db	88680	ACTTTTGGCCGCTTTCAGAAACCTTGCCATCAGCAACCAAGGCTCTTCAATTT	88621
QY	3065	CCTCGCACTCTGAGCTTACAGGTTTCCAAACGATGCTGATGAGCACTCTGCTCAACAGAT	3124
Db	88620	CCTCGCACTCTGAGCTTACAGGTTTCCAAACGATGCTGATGAGCACTCTGCTCAACAGAT	88561
QY	3125	TAAATCTTATGACCTAATATATCCAAAGGACCAAGGCTCTGATGAGCAATATCA	3184
Db	88560	----TACTTATGAGCTTAAATATATCCAAAGGACCAAGGCTCTGATGAGCAATATCA	88505
QY	3185	GCCATATCTGAGCTTATCTCATCTCCAAACCTTAAAGCACTAAGGRTTCTTGCGCT	3244
Db	88504	GCCATATCTGAGCTTATCTCATCTCCAAACCTTAAAGCACTAAGGRTTCTTGCGCT	88445
QY	3245	AAATGCTTCTGCGAATATGATTTCCCAAGTTGGCBAATATGACCAAGYCATTTATATA	3304
Db	88444	AAATGCTTCTGCGAATATGATTT--CCCAAGTTATGCGCAATATGACCAAGYCATTTATATA	88386
QY	3305	CATATATTAAGAAATCTCAAGAAACCAATATCCATTTATATAGATGATGATGATGATGAT	3364
Db	88385	CATATATTAAGAAATCTCAAGAAACCAATATCCATTTATATAGATGATGATGATGATGATGAT	88326
QY	3365	AAGTGCTTCTCAAGGCTCTTAAAGAGGCTTAAACCAAGYCCCAAGCTTAAAGTTGCC	3424
Db	88325	AAGTGCTTCTCAAGGCTCTTAAAGAGGCTTAAACCAAGYCCCAAGCTTAAAGTTGCC	88280
QY	3425	AACGAGGCAAGACTTTTSTTTATATATCAAGAAATTAACATATATGATGATGATGATGATGATGAT	3484
Db	88279	AACGAGGCAAGACTTTTSTTTATATATCAAG--AAATTAACATATATGATGATGATGATGATGATGAT	88221
QY	3485	TTACACAGTTCGATGAGGATGAGCTTCAACCTTGTGCTGATGATGATGATGATGATGATGATGAT	3544
Db	88220	TTACACAGTTCGATGAGGATGAGCTTCAACCTTGTGCTGATGATGATGATGATGATGATGATGAT	88161
QY	3545	TATGAGCAAGGCTTTCATTTGTTATGAGGATGATGATGATGATGATGATGATGATGATGATGAT	3604
Db	88160	TATGAGCAAGGCTTTCATTTGTTATGAGGATGATGATGATGATGATGATGATGATGATGATGAT	88101
QY	3605	CTGAAGAGTAAATATATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3664
Db	88100	CTGAAGAGTAAATATATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	88041
QY	3665	ATGACATCTCACTGCTTAAAGAGCTTGTGCTGTCAAGCAATGCTTAAATATCT	3724
Db	88040	ATGACATCTCACTGCTTAAAGAGCTTGTGCTGTCAAGCAATGCTTAAATATCT	87981
QY	3725	AGGCTCTATTAATCTGAAGGCTGCTGCTGATCTGCTGATCTTGAATCTTAAATCTTAAATCTTAA	3784
Db	87980	AGGCTCTATTAATCTGAAGGCTGCTGCTGATCTGCTGATCTTGAATCTTAAATCTTAAATCTTAA	87921
QY	3785	YCNATATCTTCCAGCAATATGAAGAAAGATATATATATATATATATATATATATATATATATATAT	3844
Db	87920	CCACATTTCTTCCAGCAATATGAAGAAAGATATATATATATATATATATATATATATATATATATAT	87865
QY	3845	AAACCTATGCACTCGAGGGAATTTATGATGCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGAT	3904
Db	87860	AAACCTATGCACTCGAGGGAATTTATGATGCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGAT	87802
QY	3905	TTGATATCTGATGGAATCTTCTTGTATGAAGAAAGCACTTGAATATGAGGATGATGATGATGATGATGAT	3964
Db	87801	TTGATATCTGATGGAATCTTCTTGTATGAAGAAAGCACTTGAATATGAGGATGATGATGATGATGATGAT	87742


```
OY 3965 GTGATGTAATGAAATATTGAAAGTAATCCCTCACTCCAGAACTATGCTYAGCTR 4024
|||
|||
|||
Db 87741 GTGATGTAATGAAATATTGAAAGTAATCCCTCACTCCAGAACTATGCTCACTCA 87682
|||
|||
|||
OY 4025 GCAGAACTAATAGCCCTCACTYTGKGCACATGAAATTAGGAAAGRAAAAAGGTYAATTA 4084
|||
|||
|||
Db 87681 GCAAGAACTAATAGCCCTCACTYTGKGCACATGAAATTAGGAAAGRAAAAAGGTYAATTA 87622
|||
|||
|||
OY 4085 TATAAGACTCTAATATATGCTYACCTAGTCTNTCACTGGCCATGMRGCAATATGSAAGAA 4144
|||
|||
|||
Db 87621 TATACAGACTCTAATATATGCTYACCTAGTCTNTCACTGGCCATGMRGCAATATGSAAGAA 87562
|||
|||
|||
OY 4145 AGGAAATCCCTAATCTCTGAGRGAAACACTTATCMAACATCAGAAAGCCATTAGARATTA 4204
|||
|||
|||
Db 87561 AGGAAATCCCTAATCTCTGAGRGAAACACTTATCMAACATCAGAAAGCCATTAGARATTA 87502
|||
|||
|||
OY 4205 TTATYGGCMGTACAGAAACCTTAPAGAGTGMAAGTCTTACACTGCGGGGTCACTCANA 4264
|||
|||
|||
Db 87501 TTATYGGCMGTACAGAAACCTTAPAGAGTGMAAGTCTTACACTGCGGGGTCACTCANA 87442
|||
|||
|||
OY 4265 GGAAGAAAGAAAGGAAATASAAAGAAATGCCAAGCAKATATTGAAGMAAAGAGCTGCA 4324
|||
|||
|||
Db 87441 GGAAGAAAGAAAGGAAATASAAAGAAATGCCAAGCAKATATTGAAGMAAAGAGCTGCA 87382
|||
|||
|||
OY 4325 AGGCAAGACCTCTCAATTAGAAATGCTTATTAACCTTCCCTTAGATAGGTAATCCCTTC 4384
|||
|||
|||
Db 87381 AGGCAAGACCTCTCAATTAGAAATGCTTATTAACCTTCCCTTAGATAGGTAATCCCTTC 87322
|||
|||
|||
OY 4385 CGGGAAACCAAGCCCCAGTACTCAGACGGAGAAACAGAAATGGGAAACCTCAGAG- CAG 4443
|||
|||
|||
Db 87321 CGGGAAACCAAGCCCCAGTACTCAGACGGAGAAACAGAAATGGGAAACCTCAGAG- CAG 87262
|||
|||
|||
OY 4444 TTTTCTCCCTCGGGAAGGTTAGCCACTGAAAGAAAGGAAATACTTTTGCTGCACTAT 4503
|||
|||
|||
Db 87261 TTTTCTCCCTCGGGAAGGTTAGCCACTGAAAGAAAGGAAATACTTTTGCTGCACTAT 87202
|||
|||
|||
OY 4504 CCAATGGAATTAATTAAACCTTCACTCAACCTTCACTTAGGCATCGATAGACCCA 4563
|||
|||
|||
Db 87201 CCAATGGAATTAATTAAACCTTCACTCAACCTTCACTTAGGCATCGATAGACCCA 87142
|||
|||
|||
OY 4564 TCARATGCGCAAAATCATTTATTTACTGAGACGAGCCTTTTCAAACTATCAAGCABATAT 4623
|||
|||
|||
Db 87141 TCARATGCGCAAAATCATTTATTTACTGAGACGAGCCTTTTCAAACTATCAAGCABATAT 87082
|||
|||
|||
OY 4624 CAGGGCTGTGAATGTGCGCABAAATTAATCCCTGCTATGCGCAAGCTCTTCAAG 4683
|||
|||
|||
Db 87081 CAGGGCTGTGAATGTGCGCABAAATTAATCCCTGCTATGCGCAAGCTCTTCAAG 87022
|||
|||
|||
OY 4684 ABAACAAABAACAGGCCATTACCTGABAAABACTGSCAACTGATTTTACCCACAGCCC 4743
|||
|||
|||
Db 87021 ABAACAAABAACAGGCCATTACCTGABAAABACTGSCAACTGATTTTACCCACAGCCC 86962
|||
|||
|||
OY 4744 AAACCTCAGGGATTTTCACTATCTACTGCTGGGTAAATATCTTTCAAGGGTTGGGCAAG 4803
|||
|||
|||
Db 86961 AAACCTCAGGGATTTTCACTATCTACTGCTGGGTAAATATCTTTCAAGGGTTGGGCAAG 86902
|||
|||
|||
OY 4804 GCTTCCCTGTAGAGCAGAAAAAGGCCAAGAGTAATAAGGCACTAGTTCAATGAAATA 4863
|||
|||
|||
Db 86901 GCTTCCCTGTAGAGCAGAAAAAGGCCAAGAGTAATAAGGCACTAGTTCAATGAAATA 86842
|||
|||
|||
OY 4864 ATTTCCAGATTTGGGACTTTCCCGAGGCTTACAGAGTACAAATAGCCCTGCTTTCAAGGC 4923
|||
|||
|||
Db 86841 ATTTCCAGATTTGGGACTTTCCCGAGGCTTACAGAGTACAAATAGCCCTGCTTTCAAGGC 86782
|||
|||
|||
OY 4924 ACAGTAACCAAGGAGTATCCAGGGGTTAGGTAATGATATACATTAACCTGCGCTGA 4983
|||
|||
|||
Db 86781 ACAGTAACCAAGGAGTATCCAGGGGTTAGGTAATGATATACATTAACCTGCGCTGA 86722
|||
|||
|||
OY 4984 AGGCCACAGTCTCTCAGGGAAGGTGAGAAAAATGATGAAATVACTCAAAGGCACTTTAAA 5043
|||
|||
|||
Db 86721 AGGCCACAGTCTCTCAGGGAAGGTGAGAAAAATGATGAAATVACTCAAAGGCACTTTAAA 86662
|||
|||
|||
OY 5044 AAGCAAAACCAAGAAACCACTCAATGCGCTGTCTGTGCTTATAGCTTTAAAAAGA 5103
|||
|||
|||
```



```
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68162.1"
/db_xref="GI:37544402"
/translation="MALPPIHIFLFTVLLPSTLTAPPCRCMISSSPYQEFWMORP
GNIDAPSRSLSKPTFTTAHTMPRNCYHSATLCMHANTHWKXMINPSCPGGLV
TVCMTYPTQMGSDGGVODAKREKAEVSQLTRVHGTS PYKGLDLSKHETLT
HTRLVSLFNTLTGLHEVSAONPNCWCLEPNRPVYSIIVPEOMNFSITELT
LVGLVSNLEITHTSNLCVFSNTTYTNSOCIRMTPTPOICLPSCIPEVCGTSA
YRCNLSSEMCFLSPVPMPTIYTBDLYSVYSKRNRPVILPVIAGVUGAG
TGIGITTSSTOFTYKLSOELNGMERVADSLVTDOLNSLAVALONRRRLDLTAE
RGGICLPLGECCCYVNOSGIVTEKVEIRDRIQRAEELNTPWMLLQOMEWIIP
FLGPLAIIILLFLFPCIFNLVNFVSRLEAVTLQMEPKQSKTKLYRBLDRPASP
RSDVNDIKGTPPEIRISAQPLRPNSAGSS"
9487 . 10222
```

ORIGIN
LTR

```
Query Match      85.0%; Score 6446.4; DB 9; Length 10222;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 51; Indels 53; Gaps 27;

QY 606 GATGGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGATTTCTGGAAATTGGA 665
DB 2886 GATGGGAAACGTTCCCGCAAGCAAAAACGCCCTTAAGACGATTTCTGGAAATTGGA 2945
QY 666 MCAATTTGACCCCTCAGACATTAAGAAAAGCACTATATTCCTTCTGAGTGGCCGCTG 725
DB 2946 CCAATTTGACCCCTCAGACATTAAGAAAAGCACTATATTCCTTCTGAGTGGCCGCTG 3005
QY 726 GCACTCCTGAGGAGATTAATAATTAACAACATCTTACAGCTAGACCTCTTTTGTGA 785
DB 3006 GCACTCCTGAGGAGATTAATAATTAACAACATCTTACAGCTAGACCTCTTTTGTGA 3065
QY 786 AAGGCAATGAGATGAAGTGCCTAAGTACAACTTTCTTTTCAATTAAGACAACTAC 845
DB 3066 AAGGCAATGAGATGAAGTGCCTAAGTACAACTTTCTTTTCAATTAAGACAACTAC 3125
QY 846 AATTATGTAAAAAGTGAATTTATGCCCTACAGGAGCCCTTCAGAGCTACCTCCCTATC 905
DB 3126 AATTATGTAAAAAGTGAATTTATGCCCTACAGGAGCCCTTCAGAGCTACCTCCCTATC 3185
QY 906 CCAGCATCCCGCATCTCCTTCCCAATTAAGGACCCCTTCAACCCAAATGCTCCA 965
DB 3186 CCAGCATCCCGCATCTCCTTCCCAATTAAGGACCCCTTCAACCCAAATGCTCCA 3245
QY 966 AAGGAGATGACAAAGGGTAAACAGTGAACAAAGAGTGCCTAATTTCCCAATTATGA 1025
DB 3246 AAGGAGATGACAAAGGGTAAACAGTGAACAAAGAGTGCCTAATTTCCCAATTATGA 3305
QY 1026 CCCCTCCCAAGCAGTGGGAGGAAGAATTCGGCCCAAGCAGAGTGAATGCVTTTTTY 1085
DB 3306 CCCCT -CCAAGCAGTGGGAGGAAGAATTCGGCCCAAGCAGAGTGAATGCVTTTTTY 3364
QY 1086 TCTCCGAGCTTAAGCAATAAAAAAGAGCTTAAGTAATCTCAGATAAATCTGATG 1145
DB 3365 TCTCCGAGCTTAAGCAATAAAAAAGAGCTTAAGTAATCTCAGATAAATCTGATG 3424
QY 1146 CTATATTTGATTTTACAAGGTTTGAACAATCTTTGATCTGACATGAGAGATATTA 1205
DB 3425 CTATATTTGATTTTACAAGGTTTGAACAATCTTTGATCTGACATGAGAGATATTA 3483
QY 1206 TGTCACTGCTAAATCAGACCTAACCCCAATAGAGAGTGCACACATACGCAAGCC 1265
DB 3484 TGTCACTGCTAAATCAGACCTAACCCCAATAGAGAGTGCACACATACGCAAGCC 3543
QY 1266 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGGTCAATGATANGATGACAAGAG 1325
DB 3544 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGGTCAATGATANGATGACAAGAG 3600
QY 1326 AAAAGAAATGATTTCCCAAGGCAAGGCAAGCAAGTTCCTCAATGATGAGGAG 1385
DB 3601 AAAAGAAATGATTTCCCAAGGCAAGGCAAGCAAGTTCCTCAATGATGAGGAG --- 3657
```

```
QY 1386 ACAGAAATCAGTAACATGGGAGATTTGGTGCAGACATTTGCTAATCTGTGCTASAA 1445
DB 3658 ACACAGAAATCAGAAACAT -GAGATTTGGTGTGCGACAGATTTGCTAATCTGTGCTAGAA 3716
QY 1446 GGACTAAGAAAACCTASGAGAAAATCTAYGAATTAATCTCAATGATGTCCACATACACA 1505
DB 3717 GGACTAAGAAAACCTAGAAAG -AAGTCAATGAATTAATCTCAATGATGTCCACATACACA 3775
QY 1506 GGGGAAAGGAAAGAAAATCTCACTGTGCTTTCTGAGAGAACTTAAGGAGCAATTGGAAGC 1565
DB 3776 -GGGAAAGGAAAGAAAATCTCACTGTGCTTTCTGAGAGAACTTAAGGAGCAATTGGAAGC 3834
QY 1566 GTGCCTCTCTGTCAACCTGAACCTCTTCTGAAAGGCCAATTAATCTTAAGCGTAAGTTTCA 1625
DB 3835 GTGCCTCTCTGTCAACCTGAACCTCTTCTGAAAGGCCAATTAATCTTAAGCGTAAGTTTCA 3894
QY 1626 CTCAGTCAGCTGAGACATTAAG -AAAAAATCTCAAAAAGTCTGCGTAGGCGCCGAGCAAA 1684
DB 3895 CTCAGTCAGCTGAGACATTAAGAAAAAAATCTTCAAAAAGTCTGCGTAGGCGCCGAGCAAA 3954
QY 1685 ACTTAAGAAACCTTAATGAACTTGAACCTTGCAACTTCGCTTTTATTAATAGATCAGAGAGC 1744
DB 3955 ACTTAAGAAACCTTAATGAACTTGAACCTTCGCAACTTCGCTTTTATTAATAGATCAGAGAGC 4014
QY 1745 AGGCGGAAACAGACCAACGAGGATTAATAAAAAAGGCCACGCTTTAATGTCATGACCTCAG 1804
DB 4015 AGGCGGAAACAGACCAACGAGGATTAATAAAAAAGGCCACGCTTTAATGTCATGACCTCAG 4074
QY 1805 CAACTGGAATTTGAGAGCTCTGAGAAAAGGAAAAGTGGCAATTGAATGCCATTAATAG 1864
DB 4075 CAACTGGAATTTGAGAGCTCTGAGAAAAGGAAAAGTGGCAATTGAATGCCATTAATAG 4134
QY 1865 GCTTGCTTCCAGTGGCGGTCTACAGAGCACTTTAAAAAAGATTGTCCAAGTGAAGTAA 1924
DB 4135 GCTTGCTTCCAGTGGCGGTCTACAGAGCACTTTAAAAAAGATTGTCCAAGTGAAGTAA 4194
QY 1925 CGGCCCCCTTGCTGATGAGCCCTTAATTTCAAGGGAATCACTGAGAGGCCACTGCCCCAG 1984
DB 4195 CGGCCCCCTTGCTGATGAGCCCTTAATTTCAAGGGAATCACTGAGAGGCCACTGCCCCAG 4254
QY 1985 GGAACAAGGCTCTTTTGAAGTCAAGAGCACTTAACAGATGATCCAGACAGAGCTGAGG 2044
DB 4255 GGAACAAGGCTCTTGAAGTCAAGAGCACTTAACAGATGATCCAGACAGAGCTGAGG 4314
QY 2045 TGCCTGGGCAAGCGGCATTCCTCAATGATCACTTCAAGAGCCCTGGGATGCTTGAAC 2104
DB 4315 TGCCTGGGCAAGCGGCATTCCTCAATGATCACTTCAAGAGCCCTGGGATGCTTGAAC 4374
QY 2105 ATTGAAGGCAAGGAGTGTGCTCCCTGAGACATGGTGGCGGTCTTTAATGCTTAATCTTC 2164
DB 4375 ATTGAAGGCAAGGAGTGTGCTCCCTGAGACATGGTGGCGGTCTTTAATGCTTAATCTTC 4433
QY 2165 TGTCCCGGCAACCTGTCTCTCAGATCTGCACTATTCTGAGGGGCTCCTTAAGACGGCA 2224
DB 4434 TGTCCCGGCAACCTGTCTCTCAGATCTGCACTATTCTGAGGGGCTCCTTAAGACGGCA 4491
QY 2225 GTCACTAATACCTTTTCCAGCCACTAAGTTATGAATGAGAGTGGAGCTTTATCTTTTAC 2284
DB 4492 GTCACTAATACCTTTTCCAGCCACTAAGTTATG -ACTGGGAGGCTTTATCTTTTAC 4549
QY 2285 ATGCTTTTCTAATTAATGCTTGAAGGCCCACTAATCTTGTTAAGGAGAGCACTCTAGCAA 2344
DB 4550 ATGCTTTTCTAATTAATGCTTGAAGGCCCACTAATCTTGTTAAGGAGAGCACTCTAGCAA 4609
QY 2345 AAGCAGGGGCAATTAACACTGTAACATTAAGAGAAAGAACACCCGTTTGTGTGCCCTG 2404
DB 4610 AAGCAGGGGCAATTAACACTGTAACATTAAGAGAAAGAACACCCGTTTGTGTGCCCTG 4668
QY 2405 CTTGAGGAAAGAAATTAATCTTAAGTCTGGGCAACAGAAAGCAATATGAGCAAGCCAAA 2464
DB 4669 CTTGAGGAAAGAAATTAATCTTAAGTCTGGGCAACAGAAAGCAATATGAGCAAG -CAAA 4727
```

2465 GAATGCCGCTGTTCAAGTTAACTAAAGGATTCACTTCCTTCCCTACCAAGGCA 2524
4728 GAAAGCCCGTCTCTTCAAGTTAACTAAAGGATTCACTTCCTTCCCTACCAAGGCA 4787
2525 GTACCCCTCAGACCCCAAGGCCCAACAAGATTCCAAAAGATTGTTAAGAATTAAAGC 2584
4788 GTACCCCTCAGACCCCAAGGCCCAACAAGATTCCAAAAGATTGTTAAGAATTAAAGC 4847
2585 CCAAGGCTTAGTAAACCATCATTAATCCCTGCAATATTCCTGATGGAATTGAGAGG 2644
4848 CCAAGGCTTAGTAAACCATCATTAATCCCTGCAATATTCCTGATGGAATTGAGAGG 4900
2645 CACGAAACCAGTGGACAGTGGAGGTTATGCGAAGATCTCAGGATTATAGAGAGC 2704
4901 TACGAAACCCCAACAGACAGTGA- GATTAGTCAAGATCTCAGGATTATCAAT- GAAGC 4958
2705 CGTGTCTTTTATACCCAGCTGACCTAGCCCTTAATCTGAGTTTCCAAAATACAGA 2764
4959 TGTGTCTCTTAAGCAGCTGTACTAGCCCTTAATCTGCTTCCCAATACAGA 5018
2765 GGAAGACAGTGTATTACASTCTGGAACCTTMAAGATCCTTCTTGCATCCTGTA 2824
5019 GGAAGACAGTGTATTACAGTCTGGAACCTTMAAGATCCTTCTTGCATCCTGTA 5078
2825 TCTGACTCTCAATCTTGTGCTTGAAGATCTTCAACCAACCATCTCACTAC 2884
5079 TCTGACTCTCAATCTTGTGCTTGAAGATCTTCAACCAACCATCTCACTAC 5138
2885 CTGACATCTTTCACCCCAAGGTTGAGGATAGTCCCATCTATTTGGCCAGGACTTAC 2944
5139 CTGACATCTTTCACCCCAAGGTTGAGGATAGTCCCATCTATTTGGCCAGGACTTAC 5198
2945 CCAAGACTTGAAGCARTYMTCACTAGCACTCTTCTCTCTGTAAGTGTGATTT 3004
5199 CCAAGACTTGAAGCARTYMTCACTAGCACTCTTCTCTCTGTAAGTGTGATTT 5256
3005 ACTTTTGGCTGCTTTCAGAAACCTTGTGCTATCAAGCCCAAGGCTCTTAAATT 3064
5257 ACTTTTGGCTGCTTTCAGAAACCTTGTGCTATCAAGCCCAAGGCTCTTAAATT 5316
3065 CCTGCACTCTGCTACAGGTTTCCAAACCAAGGCTCTTAAATT 3124
5317 CCTGCACTCTGCTACAGGTTTCCAAACCAAGGCTCTTAAATT 5376
3125 TAAATCTTGAAGCTTAAATTAATCAAGGCTCTTAAATT 3184
5377 TAAATCTTGAAGCTTAAATTAATCAAGGCTCTTAAATT 5432
3185 GCTTATCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3244
5433 GCTTATCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 5492
3245 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3304
5493 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3364
3305 CASTAATTAAGGAACTCAAGAAAGCAATCAATTAATTAAGTGAATCTGAAGT 3364
5522 CACTAATTAAGGAACTCAAGAAAGCAATCAATTAATTAAGTGAATCTGAAGT 5511
3365 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3424
5512 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 5567
3425 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3484
5568 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 5716
3485 TTAACAGATCCGAGGAGTGAAGTCTGCAACCTTGAAGTGAAGTGAAGTGAAGT 3544
5717 TTAACAGATCCGAGGAGTGAAGTCTGCAACCTTGAAGTGAAGTGAAGTGAAGT 5776
3545 TAGTGGCAAGGCTTGTCTTCTTATGAGGATGCTGAGGATGAGGATGAGTAT 3604

5777 TAGTGGCAAGGCTTGTCTTCTTATGAGGATGCTGAGGATGAGGATGAGTAT 5836
3605 CTGAGACATTTAAATTAATCAAGGAGAGATCTTACTGTGTGACATCTCATAGTGA 3664
5837 CTGAGACATTTAAATTAATCAAGGAGAGATCTTACTGTGTGACATCTCATAGTGA 5896
3665 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3724
5897 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 5956
3725 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3784
5957 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6016
3785 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3844
6017 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6076
3845 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3904
6077 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6135
3905 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 3964
6136 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6195
3965 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4024
6196 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6255
4025 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4084
6256 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6315
4085 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4144
6316 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6375
4145 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4204
6376 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6435
4205 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4264
6436 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6495
4265 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4324
6496 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6555
4325 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4384
6556 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6615
4385 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4443
6616 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6675
4444 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4503
6676 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6735
4504 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4563
6736 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6795
4564 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 4623
6796 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6855
4624 AATAGGTTCTGCTTATCTCTCATCTCAACCAAGGCTCTTAAATT 6883

Dh 6856 CAGGCGCTGTGAAGTGTGCGACAGAAATATCCCTGCTTATGCGCAAGCTCTTCAGG 6915
Qy 4684 ARAACAAARAAACAGGSCATTACCTGTBARABARACTGGCAACTGATTTTATCCCAACAAGGCC 4743
Dh 6916 AGAACAAAGAACAGGSCATTACCTGTGAGAGACTGGCAACTGATTTTATCCCAACAAGGCC 6975
Qy 4744 AAACCTCAGGAGTTTCACTATCTACTAGTCTGGGTARATCTTTTACAGGGTTGGCAGAG 4803
Dh 6976 AAACCTCAGGAGTTTCACTATCTACTAGTCTGGGTARATCTTTTACAGGGTTGGCAGAG 7035
Qy 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATTAAGGCACTAGTTTATTAATA 4863
Dh 7036 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATTAAGGCACTAGTTTATTAATA 7095
Qy 4864 ATTCCAGATTGCGACTTCCCGAAGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGGCC 4923
Dh 7096 ATTCCAGATTGCGACTTCCCGAAGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGGCC 7155
Qy 4924 ACAGTAAACCCAGGAGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCTGA 4983
Dh 7156 ACAGTAAACCCAGGAGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCTGA 7215
Qy 4984 AGGCCACAGTCTCAAGGAGAGTGAAGAAATGAATGAATTAATCAAAAGACATCTTAAA 5043
Dh 7216 AGGCCACAGTCTCAAGGAGAGTGAAGAAATGAATGAATTAATCAAAAGACATCTTAAA 7275
Qy 5044 AAGCAAAACCCAGGAAACCCACCTCAATAGGCTGTGCTGCTGCTTATAGCTTAAAAA 5103
Dh 7276 AAGCAAAACCCAGGAAACCCACCTCAATAGGCTGTGCTGCTTATAGCTTAAAAA 7335
Qy 5104 ATCTGCACTTTCCCAAAAAAGAGACTTAGCCCATACGAATGCTGTATGAGAGGCC 5163
Dh 7336 ATCTGCACTTTCCCAAAAAAGAGACTTAGCCCATACGAATGCTGTATGAGAGGCC 7395
Qy 5164 TTGATTAACCAATGACTTTGTGCTTGAACCCAGACAGCCACTTAGTTGAGACATCACT 5223
Dh 7396 TTGATTAACCAATGACTTTGTGCTTGAACCCAGACAGCCACTTAGTTGAGACATCACT 7455
Qy 5224 CCTTAGCAATATCAACAAGTTCTTAAAACATTAACAAGAAACCTATCCCGAAGAGG 5283
Dh 7456 CCTTAGCAATATCAACAAGTTCTTAAAACATTAACAAGAAACCTATCCCGAAGAGG 7515
Qy 5284 GAAAAAGACTATTCACCCMWGTGACATGTATTAGTCAAGTCCCTTCYCTATATCCC 5343
Dh 7516 GAAAAAGACTATTCACCCMWGTGACATGTATTAGTCAAGTCCCTTCYCTATATCCC 7575
Qy 5344 CATCCCTAGATATCTCTGGAGAGACCCCTACCCAGTCAATTTATTTAACCCCACTGCGG 5403
Dh 7576 CATCCCTAGATATCTCTGGAGAGACCCCTACCCAGTCAATTTATTTAACCCCACTGCGG 7635
Qy 5404 TTAAAGTGGCTGAGTGAAGTCTTGATATCATCACTTGAGTAAATCCCTGGAATATGCG 5463
Dh 7636 TTAAAGTGGCTGAGTGAAGTCTTGATATCATCACTTGAGTAAATCCCTGGAATATGCG 7695
Qy 5464 CAAAGGAACCTGAAAAATCAGAGACAAACCTAGCTATTCCTGTGAACCTCTAGAGATT 5523
Dh 7696 CAAAGGAACCTGAAAAATCAGAGACAAACCTAGCTATTCCTGTGAACCTCTAGAGATT 7755
Qy 5524 TGGGCTGCTCTTCAAAACAACAACAGAGAGAAAGTAACTPAAATCATTAATCCCGATG 5583
Dh 7756 TGGGCTGCTCTTCAAAACAACAACAGAGAGAAAGTAACTPAAATCATTAAT-CCCGATG 7814
Qy 5584 GSCCTCCCTTATCATATTTTCTCTKTAAGTSTTTTAAACCTSTTTGACCTCAGTCGA 5643
Dh 7815 GSCCTCCCTTATCATATTTTCTCTKTAAGTSTTTTAAACCTSTTTGACCTCAGTCGA 7874
Qy 5644 CCCCCTCAGTCCGCTGTATGACAGTATGCTCCCTTACMAAGATTTCTATGAGAGATG 5703
Dh 7875 CCCCCTCAGTCCGCTGTATGACAGTATGCTCCCTTACMAAGATTTCTATGAGAGATG 7934
Qy 5704 CAGGCTCCCGGAAATTTGATGCCCATGCTATAGAGTCTTTSTAAGGAAACCCCAACC 5763
Dh 7935 CAGGCTCCCGGAAATTTGATGCCCATGCTATAGAGTCTTTSTAAGGAAACCCCAACC 7994

Qy 5764 TTCACTGCGCAACCAATATGCCCCGCAACTGCTATCACTGTGCACTCTTTGATGCAT 5823
Dh 7995 TTCACTGCGCAACCAATATGCCCCGCAACTGCTATCACTGTGCACTCTTTGATGCAT 8054
Qy 5824 GCAATATCTCATTTATGAGACAGAAAAATGATTAATCCTAGTTGTCTGAGAGACTTGA 5883
Dh 8055 GCAATATCTCATTTATGAGACAGAAAAATGATTAATCCTAGTTGTCTGAGAGACTTGA 8114
Qy 5884 GTCACTGTCTGTGACTTACTTCAACCAACTGATATGTCTGATGCGGAGTGAAGTCAA 5943
Dh 8115 GTCACTGTCTGTGACTTACTTCAACCAACTGATATGTCTGATGCGGAGTGAAGTCAA 8174
Qy 5944 GATCAGGCAAGAGAAAAATGTAAGAAATATCCCAACTCACCGSGGTATCATGCG 6003
Dh 8175 GATCAGGCAAGAGAAAAATGTAAGAAATATCCCAACTCACCGSGGTATCATGCG 8234
Qy 6004 ACCTTAGCCCTTACAAAGACTAGATCTCTCAAAACTAGATGAAAACCTTCCGTAACCAT 6063
Dh 8235 ACCTTAGCCCTTACAAAGACTAGATCTCTCAAAACTAGATGAAAACCTTCCGTAACCAT 8294
Qy 6064 ACTGCGCTGTGAAGCTATTTAATCAACCTCACTGCGCTCAATGAGTCTCGGCCAA 6123
Dh 8295 ACTGCGCTGTGAAGCTATTTAATCAACCTCACTGCGCTCAATGAGTCTCGGCCAA 8354
Qy 6124 AACCTTACTAATGTTGGAATATGCGCCCTGCAACTTCARGCCATATGTTTCAAATGCC 6183
Dh 8355 AACCTTACTAATGTTGGAATATGCGCCCTGCAACTTCARGCCATATGTTTCAAATGCC 8414
Qy 6184 GTACTGGAACAATGGAACAACCTTCAAGCAGAAATTAACCACTTCCGTTTATGATGA 6243
Dh 8415 GTACTGGAACAATGGAACAACCTTCAAGCAGAAATTAACCACTTCCGTTTATGATGA 8474
Qy 6244 CCTCTGTTTCCATSTGGAATTAACCATATCTCAAACTCACTGTGTAAAATTTAGC 6303
Dh 8475 CCTCTGTTTCCATSTGGAATTAACCATATCTCAAACTCACTGTGTAAAATTTAGC 8354
Qy 6304 AATACTAATATGACAACAACCTCCAAATGATAGAGGGGTAACTCTCCCAACAATA 6363
Dh 8535 AATACTAATATGACAACAACCTCCAAATGATAGAGGGGTAACTCTCCCAACAATA 8594
Qy 6364 GTCTGCTACCCCTCAGAAATATTTTGTCTGTGTGATCACTCAAGCTTACGTTGTTGAT 6423
Dh 8595 GTCTGCTACCCCTCAGAAATATTTTGTCTGTGTGATCACTCAAGCTTACGTTGTTGAT 8564
Qy 6424 GGCTCTTCAGAATCTATGTGCTTCTCTCAATTTAGTGTGCCCCATGRCATCTACCT 6483
Dh 8655 GGCTCTTCAGAATCTATGTGCTTCTCTCAATTTAGTGTGCCCCATGRCATCTACCT 8714
Qy 6484 GAAACAAGTTTATACGTTATATGATATCTAAGCCCCGGAACAAAGAGTAAACCATTTCT 6543
Dh 8715 GAAACAAGTTTATACGTTATATGATATCTAAGCCCCGGAACAAAGAGTAAACCATTTCT 8774
Qy 6544 CTTTTGTTATAGAGACAGAGTGTAGGTGACATAAGTACTGGAATTTGGCGGTATCACA 6603
Dh 8775 CTTTTGTTATAGAGACAGAGTGTAGGTGACATAAGTACTGGAATTTGGCGGTATCACA 8834
Qy 6604 ACCTTACTATGATTTCTATCAAACTATCTCAAGAACTAATGAGGACATGAAACGGGTC 6663
Dh 8835 ACCTTACTATGATTTCTATCAAACTATCTCAAGAACTAATGAGGACATGAAACGGGTC 8894
Qy 6664 GCGCATCCCTGCTACCTTGAAGATCAACTTAATCCCTGACACAGATAGTCTTCRA 6723
Dh 8895 GCGCATCCCTGCTACCTTGAAGATCAACTTAATCCCTGACACAGATAGTCTTCRA 8954
Qy 6724 AATCGAAGCTTATAGACTTGTCAACCGCTGARAGAGGGGGAACCTGTTATATTTTAAAGG 6783
Dh 8955 AATCGAAGCTTATAGACTTGTCAACCGCTGARAGAGGGGGAACCTGTTATATTTTAAAGG 9014
Qy 6784 GAAAGATGCTTATATGTTATATCAATCCGAATGCTGATGAGAAAGTTTAAAGAAATT 6843
Dh 9015 GAAAGATGCTTATATGTTATATCAATCCGAATGCTGATGAGAAAGTTTAAAGAAATT 9074

QY	6844	CGAAGATCGAATATCAACGTAACAGCAAGAGAGCTTCGAAAACATCTGAGACCTCGGGGCTCTCTC	6903
Db	9075	CGAAGATCGAATATCAACGTAAGACAGAGGAGCTTCGAAAACATCTGAGACCTCGGGGCTCTCTC	9134
QY	6904	AGCGATGGAATGCCCGCGATTCCTCCCTCTTAAAGACCGCTAGACGCTAATATATTTGGCTA	6963
Db	9135	AGCGAATGATATCCCTGGAATTCCTCCCTCTTAAAGACCGCTAGACGCTAATATATTTGGCTA	9194
QY	6964	CTCTCTTTTGGAGACCTCTGATCTTTTAACTCTTGTAACTTTGTCTTCCAGATCGAA	7023
Db	9195	CTCTCTTTTGGAGACCTCTGATCTTTTAACTCTTGTAACTTTGTCTTCCAGATCGAA	9254
QY	7024	GCTGTAAACCTAACATGAGAGCCCAAGATGACGCTGCAAGACTAAGATCTACCGACACC	7083
Db	9255	GCTGTAAACCTAACATGAGAGCCCAAGATGACGCTGCAAGACTAAGATCTACCGACACC	9314
QY	7084	CTGAGACGGGCGGTAAAGCCCAAGATGATGATTAATGATCAATAAAGGACCCCTCTAG	7143
Db	9315	CTGAGACGGGCGGTAAAGCCCAAGATGATGATTAATGATCAATAAAGGACCCCTCTAG	9374
QY	7144	GAAATCTCAGCTGACAACTCTTACTACGCCCAATTAGACAGAGAAGTATGACGCGT	7203
Db	9375	GAAATCTCAGCTGACAACTCTTACTACGCCCAATTAGACAGAGAAGTATGACGCGT	9434
QY	7204	SGTGGCCAACTCTCCCAACAGCACTTAAAGTTTCTGTGATGGGGGACTGAGAGAC	7263
Db	9435	SGTGGCCAACTCTCCCAACAGCACTTAAAGTTTCTGTGATGGGGGACTGAGAGAC	9494
QY	7264	AGGATTAAGTGGATTTTCTTACGCTGATTAAGATTCCTTAAGCTTAAGCTTAAGCTTAAG	7323
Db	9495	AGGATTAAGTGGATTTTCTTACGCTGATTAAGATTCCTTAAGCTTAAGCTTAAGCTTAAG	9554
QY	7324	ACATCCACCTTTAAACAGGGGCTTGCAACTAGTACACCTGACCAATCAGAGAGCTC	7383
Db	9555	ACATCCACCTTTAAACAGGGGCTTGCAACTAGTACACCTGACCAATCAGAGAGCTC	9614
QY	7384	ACTAAATATGCTAATTAGGCAAAAGACAGAGGTTAAAGAAATAGCCATCATATTTGCTG	7443
Db	9615	ACTAAATATGCTAATTAGGCAAAAGACAGAGGTTAAAGAAATAGCCATCATATTTGCTG	9674
QY	7444	AGAGCAACAGGAGGAGCAATGATTCGGATTTAAACCAAGTTTTCAGCCGGCAACGG	7503
Db	9675	AGAGCAACAGGAGGAGCAATGATTCGGATTTAAACCAAGTTTTCAGCCGGCAACGG	9734
QY	7504	CAACCCCTTTGGGTCCCTCCCTTGTATGAGAGACTCGTTTTCATGCTATTTCACTCT	7563
Db	9735	CAACCCCTTTGGGTCCCTCCCTTGTATGAGAGACTCGTTTTCATGCTATTTCACTCT	9794
QY	7564	ATTAAATCTTGCACCTGCR 7582	
Db	9795	ATTAAATCTTGCACCTGCA 9813	
RESULT 11			
LOCUS	AY101584	10222 bp	DNA linear PRI 11-FEB-2004
DEFINITION	Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVW1 locus, allele A, complete sequence.		
ACCESSION	AY101584		
VERSION	AY101584.1	GI:37544403	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mullerkyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 10222)		
TITLE	Mallet,F., Boulton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.		
JOURNAL	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
REFERENCE	14757826		
	2 (bases 1 to 10222)		

AUTHORS Mallet, P., Bouton, O. and Oriol, G.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
FEATURES CNRS-Bioherveur, Ecole Normale Supérieure de Lyon - 46 allée
SOURCE d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1. 10222
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="79"
/isolation_source="BMC of individual 79"
/db_xref="Eaxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC"
/note="caucasian
isolated by PCR; ERV1 locus, allele A
endogenous_virus: HERV-W"
1. 780
join(248..811,2887..9813)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68163.1"
/db_xref="GI:37544404"
/translaction="MALPYRIPLFVLLPSPFLTAPPCRCMTSSSPYOEFLMORR
TNCWYRTQMGSDGGVDQARREKVKYVLSQTLVHGNSPYKGLDKLHETLR
GVNDASVRLSKGTPTPTAHTMPPNRCYSATLQCMANHTWTKMINSPEGGLGV
HTLVASI.FNTTLTGLHVBKQNTNCHICLPIAPRVYSIPVPOANNESTEINTSV
LVPLVSENEIHTSNLTCVFSNTTYTINSQCIRWTPQIVCLPFIPOGTCG
YRLCPLNSBEMCLSPFLVPMPTIYTBODLVNVIISKRNRPVILPVIAGVIGALG
TGIGGITTSTQFYKLQOELNGDMERYADSLVIOQLNSIAVWLNORLADLITAE
RGGCLCFIEGECYVYVNOISGIVTEKYEIDRIQRAEELRNTGPMGLLSOMPMILP
FLDPLAIIILLILFGPCIPLVILNVPVSRIRIAYKLMQWPMQSKTKIYRPLDRPAPB
RSDVNDIKGTPEEISAPLIRPNSGSS"
9487..10222

ORIGIN
LTR
Query Match 84.9%; Score 6440; DB 9; Length 10222;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6676; Conservative 195; Mismatches 55; Indels 53; Gaps 27;

QY	606	GATGGGAAAGCTTCCCGCAAGACAAAAAAGCCCTTAAGAGTAATCTGTGAAATTGGGA	665
DB	2886	GATGGGAAAGCTTCCCGCAAGACAAAAAAGCCCTTAAGAGTAATCTGTGAAATTGGGA	2945
QY	666	MCAATTGACCCCTTAGACACTAAGAAAGAAACGACTTATATCTTTCGAGTGGCGCTG	725
DB	2946	CCAAATTGACCCCTTAGACACTAAGAAAGAAACGACTTATATCTTTCGAGTGGCGCTG	3005
QY	726	GCACTCCGAGGGAGATATAATTAACAACATCTTACAGCTAGACCTCTTTGTAGAA	785
DB	3006	GCACTCCGAGGGAGATATAATTAACAACATCTTACAGCTAGACCTCTTTGTAGAA	3065
QY	786	AAGCAATAGAGTGAAGTGCATTAAGTACAAATCTTCTTCTTAAGAGACAATCTAC	845
DB	3066	AAGCAATAGAGTGAAGTGCATTAAGTACAAATCTTCTTCTTAAGAGACAATCTAC	3125
QY	846	AATTATGTAAGTGTGATTTATGCCCTTACAGAAACCTTCAGAGTCTACCTCCATC	905
DB	3126	AATTATGTAAGTGTGATTTATGCCCTTACAGAAACCTTCAGAGTCTACCTCCATC	3185
QY	906	CCAGCAATCCCGACTCTTCCCAATTAATTAAGAGCCCGCTTCAACCCAAATGCTCAA	965
DB	3186	CCAGCAATCCCGACTCTTCCCAATTAATTAAGAGCCCGCTTCAACCCAAATGCTCAA	3245

QY 966 AAGAGATAGACAAAGGGTAAACAGTGAACAAAGAGTGCCTAATTTCCCAATTATGA 1025
DB 3246 AAGAGATAGACAAAGGGTAAACAGTGAACAAAGAGTGCCTAATTTCCCAATTATGA 3305
QY 1026 CCCCTCCCAAGCAGTGGAGAGAAAGAAATTTGGCCCAAGAGTGCATGTGCTTTTYY 1085
DB 3306 CCCCT-CCAAAGCAGTGGAGAGAAAGAAATTCGGCCCAAGAGTGCATGTGCTTTTTC 3364
QY 1086 TCTCCAGACTTAAAGCAATTAATAACAGACTTAAATTTCCAGATAACCTGATGG 1145
DB 3365 TCTCCAGACTTAAAGCAATTAATAACAGACTTAAATTTCCAGATAACCTGATGG 3424
QY 1146 CTATATTTGTTTAAACAAGGTTAAGCAATTTGATCTGACATGAGAGATATATA 1205
DB 3425 CTATATTTGTTTAAACAAGGTTAAGCAATTTGATCTGACATGAGAGATATA-A 3483
QY 1206 TGTCACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACATACTGACGCT 1265
DB 3484 TGTCACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACATACTGACGCT 3543
QY 1266 GAGGTGTTGGCAGTCTGTGATCTCAGTCAAGTCAATGATANGATGACACAGAGG 1325
DB 3544 GAGGTGTTGGCAGTCTGTGATCTCAGTCAAGTCAATG-ATAGGATGACACAG-AGG 3600
QY 1326 AAGANAATGATTTCCCAAGGCAAGCAGACAGTCCAGTCTAACCCTCATTTGGGAGC 1385
DB 3601 AAGANAATGATTTCCCAAGGCAAGCAGACAGTCCAGTCTAACCCTCATTTGGG- 3657
QY 1386 AAGANAATCAGTAAACATGGAGATTTGGTGTGTCAGACATTTGCTAACTGTGTCTASAA 1445
DB 3658 AAGANAATCAGTAAACATGGAGATTTGGTGTGTCAGACATTTGCTAACTGTGTCTASAA 3716
QY 1446 GGAATTAAGGAAAACTASGAAGAATCTATGAATTAATCTAATGATGCCAATAACACA 1505
DB 3717 GGAATTAAGGAAAACTASGAAGAATCTATGAATTAATCTAATGATGCCAATAACACA 3775
QY 1506 GGGGAAGGGAAGAAAACTCTACTGCTTTCTGAGAGACTAAGGAGGCAATTGAGAGAC 1565
DB 3776 -GGGAAGGGAAGAAAACTCTACTGCTTTCTGAGAGACTAAGGAGGCAATTGAGAGAC 3834
QY 1566 GTGGCTCTGTGCACTGACTGCTTTCTGAGAGGCACTAATCTTAAAGCTTAATATCA 1625
DB 3835 GTGGCTCTGTGCACTGACTGCTTTCTGAGAGGCACTAATCTTAAAGCTTAATATCA 3894
QY 1626 CTCACTCAGCTCAGACATTTAG-AAAAAATTCAAAAGTCTGCGTAAAGCCGGAGCANA 1684
DB 3895 CTCACTCAGCTCAGACATTTAGAAAAAATTCAAAAGTCTGCGTAAAGCCGGAGCANA 3954
QY 1685 ACTTGAAGAACTCTATTGAACCTTGGCAACCTGCTTTTATATATAGATCAGAGAGAC 1744
DB 3955 ACTTGAAGAACTCTATTGAACCTTGGCAACCTGCTTTTATATATAGATCAGAGAGAC 4014
QY 1745 AGGCGGAACAAGAACAAAGGATTAATAAAAGGCGCACTGCTTAATGATCAGCCACAG 1804
DB 4015 AGGCGGAACAAGAACAAAGGATTAATAAAAGGCGCACTGCTTAATGATCAGCCACAG 4074
QY 1805 CAAGTGAATTTGAGAGCTCTGAGAAAAGGGAAGAAAGTGGGCAATTTGAATGCTAATAG 1864
DB 4075 CAAGTGAATTTGAGAGCTCTGAGAAAAGGGAAGAAAGTGGGCAATTTGAATGCTAATAG 4134
QY 1865 GCTTGCCTCAGTGGGCTCTACAGAGACACTTAAATAAAAGATTTGCTCAAGTGAAGTAG 1924
DB 4135 GCTTGCCTCAGTGGGCTCTACAGAGACACTTAAATAAAAGATTTGCTCAAGTGAAGTAG 4194
QY 1925 CCGGCCCTTGTGCTATGCCCCCTTATTTCAAGGGAATCACTGAGAGCCCACTGCCCCAG 1984
DB 4195 CCGGCCCTTGTGCTATGCCCCCTTATTTCAAGGGAATCACTGAGAGCCCACTGCCCCAG 4254
QY 1985 GGAAGAAAGTCTTTTGAATCAGAGACCACTAACAGATGATCCAGAGCAGAGCTGAGGG 2044
DB 4255 GGAAGAAAGTCTTTGATCAGAGACCACTAACAGATGATCCAGAGCAGAGCTGAGGG 4314
QY 2045 TGCCTGGGGCAAGCGCCATCCATGCGCATCAACCTCAAGAGCCCTGGGATGCTTGACC 2104
DB 4315 TGCCTGGGGCAAGCGCCATCCATGCGCATCAACCTCAAGAGCCCTGGGATGCTTGACC 4374
QY 2105 ATTGAGGGCCAGGAAGTGTCTCTGAGACACTGTGTGGCTCTTATGCTTAACTCTTC 2164
DB 4375 ATTGAGGGCCAGG-AGTGTGTCTGAGACACTGTGTGGGCTCTTATGCTTAACTCTTC 4433
QY 2165 TGTCCCGGCAACTGTCTCCAGACTGTGACATTTCTAGGGGGGTCCNTAAGCGGGCA 2224
DB 4434 TGTCCCGGCAACTGTCTCCAGACTGTGACATTTCTAGGGGGGTCC- TAAAGCGGGCA 4491
QY 2225 GTCACTAGATACCTTTTCCAGGCACTAAGTTATGAAGTGGAGGCTTATTTCTTTTAC 2284
DB 4492 GTCACTAGATAC- TTTCCAGGCACTAAGTTATG- ACTGGGAGCTTATTTCTTTTAC 4549
QY 2285 ATGCTTTTCTAATTTATGCTTGAAGCCCACTACCTGTTAGGAGAGACATTTAGCAA 2344
DB 4550 ATGCTTTTCTAATTTATGCTTGAAGCCCACTACCTGTTAGGAGAGACATTTAGCAA 4609
QY 2345 AAGCAGGGGCACTTATATACCTGAACATAGAGAGAGAACACCCGTTGTGTTGCCCTG 2404
DB 4610 AAGCAGGGGCACTTATATACCTGAACATAGAGAGAGAACACCCGTTGTGT- CCCCTG 4668
QY 2405 CTGAGAGAGAAATTAATCTGAAGTCTGGCAACAGAGAGCAATATGAGCAGCANA 2464
DB 4669 CTGAGAGAGAAATTAATCTGAAGTCTGGCAACAGAGAGCAATATGAGCAG- CANA 4727
QY 2465 GAATGCCGCTCTGTTCAAGTTAACTTAAGATTCACCTCTCTTCCCAAGAGCA 3524
DB 4728 GAATGCCGCTCTGTTCAAGTTAACTTAAGATTCACCTCTCTTCCCAAGAGCA 4787
QY 2525 GTACCCCTCAGACCCAGGCCCCCAAGAGATTCACAAAGATTTGTAAGACTTAAGC 2584
DB 4788 GTACCCCTCAGACCCAGGCCCCCAAGAGATTCACAAAGATTTGTAAGACTTAAGC 4847
QY 2585 CCAAGGCTTATGTAAGAACATGATTAATCCCTGACATTAATCCGATGAGATTGAGAGAG 2644
DB 4848 CCAAGGCTTATGTAAGAACATGATTAATCCCTGACATTAATCCGATGAGATTGAGAGAG 4900
QY 2645 CACAGAAACCCAGTGAAGTGGAGGTTATGTCAGAGATCTAGATTAATGATGAGAGC 2704
DB 4901 TACAGAAACCCAGTGAAGTGGAGGTTATGTCAGAGATCTAGATTAATGATGAGAGC 4958
QY 2705 GGTGTCTTTATATCCAGGCTTACCTAGCCCTTATCTGAGTTTCCCAATTCACA 2764
DB 4959 GGTGTCTTTATATCCAGGCTTACCTAGCCCTTATCTGAGTTTCCCAATTCACA 5018
QY 2765 GGAAGCAGAGTGTATTAATCACTCTGAGCCTTMAAGATGCTTCTGATCCCTGTACA 2824
DB 5019 GGAAGCAGAGTGTATTAATCACTCTGAGCCTTMAAGATGCTTCTGATCCCTGTACA 5078
QY 2825 TCTGACCTCAATTTCTGTTGCTTGAAGATTAATTTCAAAACCAACATCTCAACTGAC 2884
DB 5079 TCTGACCTCAATTTCTGTTGCTTGAAGATTAATTTCAAAACCAACATCTCAACTGAC 5138
QY 2885 CTGACCTTTTATCCCAAGGTTCAAGGATTAAGTCCCATCTAATTTGGCAGAGCAATTAGC 2944
DB 5139 CTGACCTTTTATCCCAAGGTTCAAGGATTAAGTCCCATCTAATTTGGCAGAGCAATTAGC 5198
QY 2945 CCAAGACTTGAACATTTATCTGAGACCTTGTCTCTGATGAGTGAATTT 3004
DB 5199 CCAAGACTTGAACATTTATCTGAGACCTTGTCTCTGATGAGTGAATTT 5256
QY 3005 ACTTTTGTGCTCTTCAAGAACTTGTGCTCAATCAAGCACCACAGCTCTTAAATTT 3064
DB 5257 ACTTTTGTGCTCTTCAAGAACTTGTGCTCAATCAAGCACCACAGCTCTTAAATTT 5316
QY 3065 CCTGCTAACCTGTGCTCAAGGTTTCAAAACCAAGCTCACTCTGCTCAAGCAGAGT 3124
DB 5317 CCTGCTAACCTGTGCTCAAGGTTTCAAAACCAAGCTCACTCTGCTCAAGCAGAGT 5376
QY 3125 TAAATTAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3184

QY 5344 CATCCCTAGATACATCTGGGAAGACCCCTACCTCATTTTATATACCCCAACCTGGG 5403
 DB 7576 CATCCCTAGATACATCTGGGAAGACCCCTACCTCATTTTATATACCCCAACCTGGG 7635
 QY 5404 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGAAGTCAAACTCTGGATATGCG 5463
 DB 7636 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGAAGTCAAACTCTGGATATGCG 7695
 QY 5464 CAAAGGAACCTGAAATCCAGAGACAAACCTAGCTATTCCTGGAACCTCTAGAGATT 5523
 DB 7696 CAAAGGAACCTGAAATCCAGAGACAAACCTAGCTATTCCTGGAACCTCTAGAGATT 7755
 QY 5524 TGGCGCTGCTCTTCAACCAACCAAGAGAAAGTAACTAAATCAATAATCCCCATG 5583
 DB 7756 TGGCGCTGCTCTTCAACCAACCAAGAGAAAGTAACTAAATCAATAATCCCCATG 7814
 QY 5584 GSCCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCTSTTTCACTCTCATGCA 5643
 DB 7815 GSCCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCTSTTTCACTCTCATGCA 7874
 QY 5644 CCCCCCTCAATGCGCTGATAGACAGAGCTCCCTTACCMAGAGTTTCTATGAGAAATG 5703
 DB 7875 CCCCCCTCAATGCGCTGATAGACAGAGCTCCCTTACCMAGAGTTTCTATGAGAAATG 7934
 QY 5704 CAGCGTCCGGAATATTTGATGCCCCATCGTATAGAGTCTTTSTAAGGGAACCCCAAC 5763
 DB 7935 CAGCGTCCGGAATATTTGATGCCCCATCGTATAGAGTCTTTSTAAGGGAACCCCAAC 7994
 QY 5764 TTCACTGCCCCAACCATATGCCCCGCACTGCTATCATCTGCACTTTTGCATGCAAT 5823
 DB 7995 TTCACTGCCCCAACCATATGCCCCGCACTGCTATCATCTGCACTTTTGCATGCAAT 8054
 QY 5824 GGAATATCATATTTAGGACAGAGAAATGATTAATCTTATGTTGCTCTGAGAGACTTGA 5883
 DB 8055 GGAATATCATATTTAGGACAGAGAAATGATTAATCTTATGTTGCTCTGAGAGACTTGA 8114
 QY 5884 GTCACTGTCTGTGGACTTACTTCAACCAACCTGATGCTGATGGGGTGGAGTTTCA 5943
 DB 8115 GTCACTGTCTGTGGACTTACTTCAACCAACCTGATGCTGATGGGGTGGAGTTTCA 8174
 QY 5944 GATCAGGCAAGAGAAATCATGTAAAGATTAATCTTCCCACTCAACGCGGTACATGCG 6003
 DB 8175 GATCAGGCAAGAGAAATCATGTAAAGATTAATCTTCCCACTCAACGCGGTACATGCG 8234
 QY 6004 ACCCTAGCCCCCTACAAAGAGACTAGATCTCTCAAACTACAGAAACCTCCGTACCCAT 6063
 DB 8235 ACCCTAGCCCCCTACAAAGAGACTAGATCTCTCAAACTACAGAAACCTCCGTACCCAT 8294
 QY 6064 ACTGCGCTGGTAAGCCTATTTAATACCACTCTCACTGGGCTCATGAGGCTCGAGCCAA 6123
 DB 8295 ACTGCGCTGGTAAGCCTATTTAATACCACTCTCACTGGGCTCATGAGGCTCGAGCCAA 8354
 QY 6124 AACCTTACTAATCTGTTGATATGCTCCCTGGAACCTTCAAGCCATATGTTCAATCCCT 6183
 DB 8355 AACCTTACTAATCTGTTGATATGCTCCCTGGAACCTTCAAGCCATATGTTCAATCCCT 8414
 QY 6184 GATCCTGAACATGGAACAACTTCAAGAGAAATTAACAACAACCTTCCGTTTATGAGGA 6243
 DB 8415 GATCCTGAACATGGAACAACTTCAAGAGAAATTAACAACAACCTTCCGTTTATGAGGA 8474
 QY 6244 CCTCTTGTGTTTCAATSTGGAATTAACCCATACCTCAAACTCAACCTGTGTTAAATTTAC 6303
 DB 8475 CCTCTTGTGTTTCAATSTGGAATTAACCCATACCTCAAACTCAACCTGTGTTAAATTTAC 8534
 QY 6304 AATATCATATACAAACCAACCTCCATAGCATCAAGTGGGTAATCTTCCCAACAATA 6363
 DB 8535 AATATCATATACAAACCAACCTCCATAGCATCAAGTGGGTAATCTTCCCAACAATA 8594
 QY 6364 GTTCGCTCAACCTCGAGGAATTTTTTGTCTGTGGAATCTCAAGCCTATGTTGTTGAAT 6423
 DB 8595 GTTCGCTCAACCTCGAGGAATTTTTTGTCTGTGGAATCTCAAGCCTATGTTGTTGAAT 8654

QY 6424 GGCTCTTCAAGATCTATGCTCTTCTCTCATTTTATGAGTCCCCCYATGRCATCTACAT 6483
 DB 8655 GGCTCTTCAAGATCTATGCTCTTCTCTCATTTTATGAGTCCCCCYATGRCATCTACAT 8714
 QY 6484 GAACAAGATTTATACAGTTATGATCATATCTAAACCCCGCAACAAAGAGATACCATCTT 6543
 DB 8715 GAACAAGATTTATACAGTTATGATCATATCTAAACCCCGCAACAAAGAGATACCATCTT 8774
 QY 6544 CCTTTGTTATAGAGACAGAGTGTAGTGCATAGCTATGCTGCAATTTGCGGTATCACA 6603
 DB 8775 CCTTTGTTATAGAGACAGAGTGTAGTGCATAGCTGCAATTTGCGGTATCACA 8834
 QY 6604 ACCTTACTCAAGTTTACTACAAACATATCTCAAGAACTTAAATGGGACATGGAACGGGTC 6663
 DB 8835 ACCTTACTCAAGTTTACTACAAACATATCTCAAGAACTTAAATGGGACATGGAACGGGTC 8894
 QY 6664 GCGCACTCCGTGATCACTTGCAGATCACTTAACTCCCTAGCAGCATGATGCTTGA 6723
 DB 8895 GCGCACTCCGTGATCACTTGCAGATCACTTAACTCCCTAGCAGCATGATGCTTGA 8954
 QY 6724 AATCGAAGCTTTTATGACTTGTAAACCGCTGABAGAGGGGAACTTGTATTTTAAAGG 6783
 DB 8955 AATCGAAGCTTTTATGACTTGTAAACCGCTGABAGAGGGGAACTTGTATTTTAAAGG 9014
 QY 6784 GAAAGATGCTTATATATGTTATATCAATCCGGAATGTCACAGAGAAAGTTTAAAGAAAT 6843
 DB 9015 GAAAGATGCTTATATATGTTATATCAATCCGGAATGTCACAGAGAAAGTTTAAAGAAAT 9074
 QY 6844 CSAAGATCGAATCAACGTAAGACAGAGAGTTCGAAACATCGSACCTTGGGCTCTCTC 6903
 DB 9075 CGAAGATCGAATCAACGTAAGACAGAGAGTTCGAAACATCGSACCTTGGGCTCTCTC 9134
 QY 6904 ACCGATGAGTATCCCTGGAATTTCCCTTTCTTAAAGACCTTACAGCTATATATGCTA 6963
 DB 9135 ACCGATGAGTATCCCTGGAATTTCCCTTTCTTAAAGACCTTACAGCTATATATGCTA 9194
 QY 6964 CTCCTCTTTGAGACCTGATCTTTTACCTCCCTGTTAACTTTGCTCTTCAAGAAATCGAA 7023
 DB 9195 CTCCTCTTTGAGACCTGATCTTTTACCTCCCTGTTAACTTTGCTCTTCAAGAAATCGAA 9254
 QY 7024 GCTGTAAACCTAACATGAGAGCCCAAGATGACGTCACAGACTTACCGCAGACCC 7083
 DB 9255 GCTGTAAACCTAACATGAGAGCCCAAGATGACGTCACAGACTTACCGCAGACCC 9314
 QY 7084 CTGACCGGCTGTATGAGCCCAAGATGATATGATATGACATCAAGAGCACCTCTCTGAG 7143
 DB 9315 CTGACCGGCTGTATGAGCCCAAGATGATATGATATGACATCAAGAGCACCTCTCTGAG 9374
 QY 7144 GAAATCTCAGCTGACACACTCTACTACGCCCCAATTCAGAGAGAGAGCTTATGAGCGGT 7203
 DB 9375 GAAATCTCAGCTGACACACTCTACTACGCCCCAATTCAGAGAGAGAGCTTATGAGCGGT 9434
 QY 7204 SGTGCGCAACCTCCCAACAGCACTTATGAGTTTCTGTTGATGAGGGAATGAGAGAC 7263
 DB 9435 SGTGCGCAACCTCCCAACAGCACTTATGAGTTTCTGTTGATGAGGGAATGAGAGAC 9494
 QY 7264 AGGACTACCTGAGTTTCTTAAAGCTATTAAGAACTCTTAAAGCTTAAAGCTTAAAGCT 7323
 DB 9495 AGGACTACCTGAGTTTCTTAAAGCTATTAAGAACTCTTAAAGCTTAAAGCTTAAAGCT 9554
 QY 7324 ACATCACAACCTTAAACAGAGGGGCTTGCACCTTATGATGACACTGACATGAGAGCTC 7383
 DB 9555 ACATCACAACCTTAAACAGAGGGGCTTGCACCTTATGATGACACTGACATGAGAGCTC 9614
 QY 7384 ACTAAATGCTAATTTAGCAAGACAGAGGTTAAAGAAATGACCAATCATATTTATGCTG 7443
 DB 9615 ACTAAATGCTAATTTAGCAAGACAGAGGTTAAAGAAATGACCAATCATATTTATGCTG 9674
 QY 7444 AGAGCAGAGAGAGAGAGAGATATGATCGGGAATTAATCCCAAGTATTTGAGCGGCAAGG 7503
 DB 9675 AGAGCAGAGAGAGAGAGAGATATGATCGGGAATTAATCCCAAGTATTTGAGCGGCAAGG 9734
 QY 7504 CAACCCCTTTGGGTCCTCTCTTGTATGAGAGCTGTGTTTCAATGCTATTTCACTCT 7563

Db 9735 CAACCCCTTGGTGCCCTCCCTTTATGGAGGCTCTGTTTCATGATTTCACTCT 9794
QY 7564 ATTAAATCTTGCACTGCR 7582
Db 9795 ATTAATCTTGCAACTGCA 9813

RESULT 12
BD221808 10499 bp DNA linear PAT 17-JUL-2003
LOCUS BD221808
DEFINITION Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses.
ACCESSION BD221808.1 GI:33031578
VERSION BD221808
KEYWORDS JP 2002518051-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10499)
AUTHORS Allet, P.M., Perin, J.P. and Rieger, F.
TITLE Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses
JOURNAL Patent: JP 2002518051-A/3 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002518051-A/3
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PR 23-JUN-1998 FR 98/07920
PI PARRICK M ALLET, JEAN PIERRE PERIN, FRANCOIS RIEGER, PC
C12N15/00, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC
A61P25/00,
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..10499
FT 1..10499 /organism="Homo sapiens (human)"
FEATURES
source 1..10499
location/Qualifiers
1..10499 /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 84.9%; Score 6436; DB 6; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;
QY 606 GATGGGAAAGCTTCCCGCAAGCAAAAAAGCCCTTAAGAGTATCTGAAATTTGGGA 665
Db 2953 GATGGGAAAGCTTCCCGCAAGCAAAAAAGCCCTTAAGAGTATCTGAAATTTGGGA 3012
QY 666 MCAAATTGACCTCTAGACACTAAGAAAGAAAGCACTTATTTCTTTCGCACTGCCGCTG 725
Db 3013 CCAATTGACCTCTAGACACTAAGAAAGAAAGCACTTATTTCTTTCGCACTGCCGCTG 3072
QY 726 GCACTCCTGAGGGAGTAAATTAATTAACACATCTTACAGCTGACTCTTTTGTGAA 785
Db 3073 GCACTCCTGAGGGAGTAAATTAATTAACACATCTTACAGCTGACTCTTTTGTGAA 3132
QY 786 AAGGCAATGAGTGAAGTGCATAGTACAACTTTCTTTTCAATTAGAGACAATCAC 845
Db 3133 AAGGCAATGAGTGAAGTGCATAGTACAACTTTCTTTTCAATTAGAGACAATCAC 3192
QY 846 AATTATGTAAGAAAGTGTGATTTATGCTTACAGAAAGCTTCAAGTCTACTCTCCATC 905
Db 3193 AATTATGTAAGAAAGTGTGATTTATGCTTACAGAAAGCTTCAAGTCTACTCTCCATC 3252
QY 906 CAGCAATCCCGACTCTTCCCGCAATTAATTAAGAGCCCGCTTAACCAATAGTCCA 965

Db 3253 CAGCAATCCCGACTCTTCCCGCAATTAATTAAGAGCCCGCTTAACCAATAGTCCA 3312
QY 966 AAGGAGATGACAAAGGGTAAACAAGTAAACAAAGAGTGCCTAATTTCCCAATTATGA 1025
Db 3313 AAGGAGATGACAAAGGGTAAACAAGTAAACAAAGAGTGCCTAATTTCCCAATTATGA 3372
QY 1026 CCCCTCCCAAGCAGTGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTYY 1085
Db 3373 CCCCT-CCAAGCAGTGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTTC 3431
QY 1086 TCTCCCACTTAAAGCAATTAATAACAAGTAAATTTCTCAGTATATCTGATG 1145
Db 3432 TCTCCCACTTAAAGCAATTAATAACAAGTAAATTTCTCAGTATATCTGATG 3491
QY 1146 CTATATGATGTTTAAAGAGTGAAGTGAATTTCTTGAATGAGTGAAGTGAATTA 1205
Db 3492 CTATATGATGTTTAAAGAGTGAAGTGAATTTCTTGAATGAGTGAAGTGAATTA 3550
QY 1206 TGTCACTGCTAATAGACACTAACCCCAATGAGAGAGTGCACATTAATGACGCT 1265
Db 3551 TGTCACTGCTAATAGACACTAACCCCAATGAGAGAGTGCACATTAATGACGCT 3610
QY 1266 GAGGTTTGGGATCTCTGATTTCTCAGTCAAGTCAATGATGATGATGACACAGAG 1325
Db 3611 GAGGTTTGGGATCTCTGATTTCTCAGTCAAGTCAATGATGATGATGACACAGAG 3667
QY 1326 AAGAGAAATGATTTCCCAAGCCAGCAGCAGGCTCCAGTCTAAGCCCTCATTTGGAG 1385
Db 3668 AAGAGAAATGATTTCCCAAGCCAGCAGCAGGCTCCAGTCTAAGCCCTCATTTGGAG 3724
QY 1386 ACAGAAATGATTAACATGAGAGTGGAGTGGTCTGAGCAATTTGCTAATCTGTGCTA 1445
Db 3725 ACAGAAATGATTAACATGAGAGTGGAGTGGTCTGAGCAATTTGCTAATCTGTGCTA 3783
QY 1446 GAGCTAAGGAAACTAAGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAG 1505
Db 3784 GAGCTAAGGAAACTAAGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAG 3842
QY 1506 GGGGAAAGGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAG 1565
Db 3843 GGGGAAAGGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAG 3901
QY 1566 GTGCTCTGTGTCACCTGACTCTTCTGAAGGCCAATTAATCTTAAGCGTAAATTCA 1625
Db 3902 GTGCTCTGTGTCACCTGACTCTTCTGAAGGCCAATTAATCTTAAGCGTAAATTCA 3961
QY 1626 CTCAGTCAAGTGCAGACATTAG-AAAAAACTTAAAGTGTGCGGCGCCGAGAGCAA 1684
Db 3962 CTCAGTCAAGTGCAGACATTAG-AAAAAACTTAAAGTGTGCGGCGCCGAGAGCAA 4021
QY 1685 ACTTAAAGAACTTAAGTGAAGTGGCAACTGCGGTTTAAATTAAGATCAGAGAGAGC 1744
Db 4022 ACTTAAAGAACTTAAGTGAAGTGGCAACTGCGGTTTAAATTAAGATCAGAGAGAGC 4081
QY 1745 AGCGGAAAGAGCAAAAGGAAATTAATAAGGCAACGCTTTAATGATCATGACCTCAGG 1804
Db 4082 AGCGGAAAGAGCAAAAGGAAATTAATAAGGCAACGCTTTAATGATCATGACCTCAGG 4141
QY 1805 CAAAGTGAATTTGAAGGCTTGGAAAGGAAAGGCAATTTAAGTGTCCAGTGAAGTAA 1864
Db 4142 CAAAGTGAATTTGAAGGCTTGGAAAGGAAAGGCAATTTAAGTGTCCAGTGAAGTAA 4201
QY 1865 GCTTGTCTTCAAGGCGGCTCAAGGCACTTTAATAAGATTTCCAGTGAAGTAA 1924
Db 4202 GCTTGTCTTCAAGGCGGCTCAAGGCACTTTAATAAGATTTCCAGTGAAGTAA 4261
QY 1925 CCGCCCTTGTGTCATGCGCCCTTATTCAAGGAGTCACTGAAGGCGCCATGCGCCAGG 1984
Db 4262 CCGCCCTTGTGTCATGCGCCCTTATTCAAGGAGTCACTGAAGGCGCCATGCGCCAGG 4321
QY 1985 GAGCAAGAGTCTTTGATGAGAGGCACTAACCAATGATCCAGAGCAGAGATGAGG 2044

Db 4322 GGAAGAGTCTCTGAGTCAGAAAGCCATTAACAGATGATCCAGCAGCAGAGCTGAGG 4381
Qy 2045 TGCTGGGGGAGAGCGCCATCCATGACATCACCCTCAGAGGCCGGGATGATGTTGACC 2104
Db 4382 TGCTGGGGGAGAGCGCCATCCATGACATCACCCTCAGAGGCCGGGATGATGTTGACC 4441
Qy 2105 ATTGAAGGCGGAGAGGTTGCTCTGAGACATGTCGGTCTTCTTAAGTCTTAAGTCTTC 2164
Db 4442 ATTGAAGGCGGAGAGGTTGCTCTGAGACATGTCGGTCTTCTTAAGTCTTAAGTCTTC 4500
Qy 2165 TGTCCCGGCAAACTGTCCTCCCAATCTGTACTATTTCTGAGGGGTCCTTAAGCGGCA 2224
Db 4501 TGTCCCGGCAAACTGTCCTCCCAATCTGTACTATTTCTGAGGGGTCCTTAAGCGGCA 4558
Qy 2225 GTCACTAGATCTTCTTCCAGCCACTAAGTTATGAAGTGGAGGCTTATCTTTTCC 2284
Db 4559 GTCACTAGATCTTCTTCCAGCCACTAAGTTATG-ACGTGGAGCTTATCTTTTCC 4616
Qy 2285 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTGAGGAGACATTCTAGCA 2344
Db 4617 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTGAGGAGACATTCTAGCA 4676
Qy 2345 AAGCAGGGGCGCTTATACCTGTAACATAGAGAAAGAAACCCGTTTGTGTCCTG 2404
Db 4677 AAGCAGGGGCGCTTATACCTGTAACATAGAGAAAGAAACCCGTTTGTGTCCTG 4735
Qy 2405 CTGAGAGAGAAATTAATCTGTAAGTCTGGGCAACAGAAAGCAATATGAGAGCAAA 2464
Db 4736 CTGAGAGAGAAATTAATCTGTAAGTCTGGGCAACAGAAAGCAATATGAGAGCAAA 4794
Qy 2465 GAATCCCGCTCTGTTCAAGTTAACTAAAGATTCACCTTCTTCCCTACCAAGGCA 2524
Db 4795 GAATCCCGCTCTGTTCAAGTTAACTAAAGATTCACCTTCTTCCCTACCAAGGCA 4854
Qy 2525 GTACCCCTCTCAAGCCCAAGGCCCAAGAGATTCCAAAGATTGTTAAGACTTAAAGC 2584
Db 4855 GTACCCCTCTCAAGCCCAAGGCCCAAGAGATTCCAAAGATTGTTAAGACTTAAAGC 4914
Qy 2585 CCAAGGCTTAAATTAACATGACATACCTGCGACATTAATTCGATGATGAGAGAG 2644
Db 4915 CCAAGGCTTAAATTAACATGACATACCTGCGACATTAATTCGATGATGAGAGAG 4967
Qy 2645 CACAGAAACCAAGTGAAGAGGTTAGTGAAGATCTGAGATTAATCAATGAGAGC 2704
Db 4968 TACAGAAACCAAGTGAAGAGGTTAGTGAAGATCTGAGATTAATCAAT- GAGGC 5025
Qy 2705 CGTGTCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTCCCAATATCAGA 2764
Db 5026 TGTGTCTCTTATAGCCAGCTGTACTAGCCCTTATCTGTGMYTCCCAATATCAGA 5085
Qy 2765 GGAAGCAGAGTGGTTACASTCTGAGACCTTMAAGATGCTTCTTGTGATCCCTGTACA 2824
Db 5086 GGAAGCAGAGTGGTTACASTCTGAGACCTTMAAGATGCTTCTTGTGATCCCTGTACA 5145
Qy 2825 TCTCTACTCAATCTTGTGTTGCTTTGAGATATCTTCAAAACCCACATCTCAACTAC 2884
Db 5146 TCTCTACTCAATCTTGTGTTGCTTTGAGATATCTTCAAAACCCACATCTCAACTAC 5205
Qy 2885 CTGACTTTTATACCCCAAGGTTCAAGGATAGYCCCATCTATTTGGCCAGGACTTAC 2944
Db 5206 CTGACTTTTATACCCCAAGGTTCAAGGATAGYCCCATCTATTTGGCCAGGACTTAC 5265
Qy 2945 CCAAGACTTGAAGCTATCTATCTGAGACCTTGTCTTGTGATGAGGATGATTT 3004
Db 5266 CCAAGACTTGAAGCTATCTATCTGAGACCTTGTCTTGTGATGAGGATGATTT 5323
Qy 3005 ACTTTTGCAGCTTGAAGACCTTGTGATGAGGCAAGCAGCAGCTCTTAAATTT 3064
Db 5324 ACTTTTGCAGCTTGAAGACCTTGTGATGAGGCAAGCAGCAGCTCTTCAATTT 5383
Qy 3065 CTTGCTTACTGCTGCTACAGGTTTCCAAACSAARAGCTCARCTTGTCTACAGCAGT 3124
Db 5384 CTTGCTTACTGCTGCTACAGGTTTCCAAACSAARAGCTCARCTTGTCTACAGCAGT 5443

Qy 3125 TAAATCTTAGGCTAARATTAATCCAAAGGCAACAGGCCCTCAGTGAAGAAATVATCCA 3184
Db 5444 ----TACTTAGGCTAARATTAATCCAAAGGCAACAGGCCCTCAGTGAAGAAATVATCCA 5499
Qy 3185 GCTTACTAGGCTTATCCCTCATCTGCAAAACCTTAAAGCACTAAGGRRATTCCTGGCCT 3244
Db 5500 GCTTACTAGGCTTATCCCTCATCCCAAAACCTTAAAGCACTAAGGRRATTCCTGGCCT 5559
Qy 3245 AAYAGGTTCTCCGAAATATGATTCCTCAGGTTGCGCAATATGACAGGCTTAAATTA 3304
Db 5560 AAYAGGTTCTCCGAAATATGATTCCTCAGGTTGCGCAATATGACAGGCTTAAATTA 5618
Qy 3305 CASTAATTAAGGAACTCAGAAAGCAATACCATTATTAATGATGATGATGATGATGAT 3364
Db 5619 CASTAATTAAGGAACTCAGAAAGCAATACCATTATTAATGATGATGATGATGATGATG 5678
Qy 3365 AAGTGGCTTCCAGGCCCTTAAAGAGGCTTAAACCCAGYCCAGTGTAAAGTTGTC 3424
Db 5679 AAGTGGCTTCCAGGCCCTTAAAGAGGCTTAAACCCAGYCCAGTGTAAAGTTGTC 5724
Qy 3425 AACAGGCAAGACTTTTATTATATGATGATGATGATGATGATGATGATGATGATGATG 3484
Db 5725 AACAGGCAAGACTTTTATTATATGATGATGATGATGATGATGATGATGATGATGATG 5783
Qy 3485 TTACAGATCCAGAGGAGYAGCTTGAACCTGTCACCTGTCACCTGTCACCTGTCACCT 3544
Db 5784 TTACAGATCCAGAGGAGYAGCTTGAACCTGTCACCTGTCACCTGTCACCTGTCACCT 5843
Qy 3545 TAGTGGCAAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3604
Db 5844 TAGTGGCAAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5903
Qy 3605 CTGAAGCACTTAAATTAATCAAGGAGATCTTACTGTGTGACATCTCATGAKGTGA 3664
Db 5904 CTGAAGCACTTAAATTAATCAAGGAGATCTTACTGTGTGACATCTCATGAKGTGA 5963
Qy 3665 AYRGATATCTACCTGCTTAAAGAGACTTGTGCTGTCAGACAAAGTGTACTTAATATTC 3724
Db 5964 AYRGATATCTACCTGCTTAAAGAGACTTGTGCTGTCAGACAAAGTGTACTTAATATTC 6023
Qy 3725 AAGCTTATTAATGAAAGGCAAGTCTGCACTGTGCACTTGTGCACTTGTGCACTTGTG 3784
Db 6024 AAGCTTATTAATGAAAGGCAAGTCTGCACTGTGCACTTGTGCACTTGTGCACTTGTG 6083
Qy 3785 YCNCAATTTCTTCCAGACATGAAGAAAGATTAATTAATTAATTAATTAATTAAT 3844
Db 6084 YCNCAATTTCTTCCAGACATGAAGAAAGATTAATTAATTAATTAATTAATTAAT 6143
Qy 3845 AAACCTATGCACTGAGAGGCACTTATGATGATGATGATGATGATGATGATGATGATG 3904
Db 6144 AAACCTATGCACTGAGAGGCACTTATGATGATGATGATGATGATGATGATGATGATG 6202
Qy 3905 TTTGATATGATGAAAGTTCTTGTGATGAAAGAAAGCTTGAAGAGGATGATGATGATG 3964
Db 6203 TTTGATATGATGAAAGTTCTTGTGATGAAAGAAAGCTTGAAGAGGATGATGATGATG 6262
Qy 3965 GTCACTGATTAATGAAATTAATGAAAGTATCCCTCATCTCAGAGAACTATGCTTAAG 4024
Db 6263 GTCACTGATTAATGAAATTAATGAAAGTATCCCTCATCTCAGAGAACTATGCTTAAG 6322
Qy 4025 GCAGAACTTAATGAGCTCAATYKGGGCACTAGAAATTAAGAGAAAGGATTAATTA 4084
Db 6323 GCAGAACTTAATGAGCTCAATYKGGGCACTAGAAATTAAGAGAAAGGATTAATTA 6382
Qy 4085 TATAAGACTTATATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4144
Db 6383 TATAAGACTTATATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 6442
Qy 4145 AGGGAATTTCTTAACTTCTGAGAGAACTATGATGATGATGATGATGATGATGATGAT 4204
Db 6443 AGGGAATTTCTTAACTTCTGAGAGAACTATGATGATGATGATGATGATGATGATGAT 6502

QY	4205	TTATTCGCMGTACAGAAACCTTAPAGAGGTGMACTCTTACATCCTCCYGGGGTCATCAAAAA	4264
Dp	6503	TTATTCGCTGTACAGAAAACCTTAAAGAGGTGGCAGCTTACATCCTCCGGGGTCATCAAAAA	6562
QY	4265	GGAAGAATAAGGAAAAATASAAGRGAAYTGCACAGACAKATATTGAGGMAAAAGAGCTGCA	4324
Dp	6563	GGAAGAATAAGGAAAAATGAAAGAGAACTGCCACAGCATATTGAAAGCCAAAAGAGCTGCA	6622
QY	4325	AGGCGAGACCTCTCCATTGAAATGCTTATTAACCTTCCCTTAGATAGGGTAACTCCTTC	4384
Dp	6623	AGGCGAGACCTCTCCATTGAAATGCTTATTAACCAACCCCTTAGATAGGGTAACTCCTTC	6682
QY	4385	CGGGAATACCAAGCCCCAGTACTCCAGAGAGAAACAGAAATGGGGAAACCTCACAGAG-CAG	4443
Dp	6683	CGGGAATACCAAGCCCCAGTACTCCAGAGAGAAACAGAAATGGGGAAACCTCACAGAGACAG	6742
QY	4444	TTTTCTCCCTCGGAGCGGTTAGCCATCTGAAGAAAGGAAAAATACTTTTGCTGCACATAT	4503
Dp	6743	TTTTCTCCCTCGGAGCGGTTAGCCATCTGAAGAAAGGAAAAATACTTTTGCTGCACATAT	6802
QY	4504	CCAAATGNAATTACTTAAAAACCTTTCATCAACCTTTCATTAGGCAATCGATAGCACCA	4563
Dp	6803	CCAAATGNAATTACTTAAAAACCTTTCATCAACCTTTCATTAGGCAATCGATAGCACCA	6862
QY	4564	TCGAAATGCGCAATCATTTATTTATCTGGACCGAGGCTTTCAAAACATACAGCABATATKT	4623
Dp	6863	TCGAAATGCGCAATCATTTATTTATCTGGACCGAGGCTTTCAAAACATACAGCABATATGT	6922
QY	4624	CAGGGCCGTGAATGACATGACABARAAATATATCCCTGCCTATGCGCCAAAGCTCCTGAG	4683
Dp	6923	CAGGGCCGTGAATGACATGACABARAAATATATCCCTGCCTATGCGCCAAAGCTCCTGAG	6982
QY	4684	ARAAACAAARAAACAGGCCATTAACCTGABARAAACTGCGAACTGATTTTACCAACAGCC	4743
Dp	6983	ARAAACAAARAAACAGGCCATTAACCTGABARAAACTGCGAACTGATTTTACCAACAGCC	7042
QY	4744	AAACCTCAGGATTTCAATATCTACTATCTGGGATATATCTTTCAAGGGTTGGGACBAG	4803
Dp	7043	AAACCTCAGGATTTCAATATCTACTATCTGGGATATATCTTTCAAGGGTTGGGACBAG	7102
QY	4804	GCCTTCCCCCTGTAGGACAGAAAAAGCCCCAAGAGTAATTAAGGCACTAGTTCAATGAATA	4863
Dp	7103	GCCTTCCCCCTGTAGGACAGAAAAAGCCCCAAGAGTAATTAAGGCACTAGTTCAATGAATA	7162
QY	4864	ATTCCCAATTCGACATTTCCCGGAGCTTACAGATGACAAATAGCCCTTCCTTCCAGGCC	4923
Dp	7163	ATTCCCAATTCGACATTTCCCGGAGCTTACAGATGACAAATAGCCCTTCCTTCCAGGCC	7222
QY	4924	ACAGTAAACCCAGGAGATATCCAGGCGTTAGGTATACATATCACTTACATGCGGCCGGA	4983
Dp	7223	ACAGTAAACCCAGGAGATATCCAGGCGTTAGGTATACATATCACTTACATGCGGCCGGA	7282
QY	4984	AGGCCACAGTCTCTCAGGGAGGTTCAGAAAAATGAATGAAYACTCAAGAGCACTCTAAAAA	5043
Dp	7283	AGGCCACAGTCTCTCAGGGAGGTTCAGAAAAATGAATGAAYACTCAAGAGCACTCTAAAAA	7342
QY	5044	AAGCAAAACCCAGAAAAACCACTCCATATGCGCTGTCTGTTGCTTATAGCCTTAAAAAGA	5103
Dp	7343	AAGCAAAACCCAGAAAAACCACTCCATATGCGCTGTCTGTTGCTTATAGCCTTAAAAAGA	7402
QY	5104	ATCTGCAACTTTCCCAAAAAGCAGACTTATAGCCCATACGAAATGCTGTATGAAAGGCC	5163
Dp	7403	ATCTGCAACTTTCCCAAAAAGCAGACTTATAGCCCATACGAAATGCTGTATGAAAGGCC	7462
QY	5164	TTCAATPACCAATGACCTTGTGCTGACCCAAAGCAGCCAACTTAGTGGACATCATCCT	5223
Dp	7463	TTCAATPACCAATGACCTTGTGCTGACCCAAAGCAGCCAACTTAGTGGACATCATCCT	7522
QY	5224	CCTTAGCCAAATATCAACAAGTTCTTAAAACAATTACAGAAACCTATCCTCGAGAAAGG	5283
Dp	7523	CCTTAGCCAAATATCAACAAGTTCTTAAAACAATTACAGAAACCTATCCTCGAGAAAGG	7582
QY	5284	GAAAAAGACTATTCACCCCMGTGACATGTATTAGTCAAGTCCCTTCCTCTAATTCC	5344

Db	7583	GAAGAAACATATTCACCTCTTGACATGGATTAATGACAGTCCCTCCCTCTAATTC	7642
OY	5344	CATCCCTAAGTATCATCTCTGGAAAGAACCCCTACCCAGTCATTTATATACCCCACTGGG	5403
Db	7643	CATCCCTAAGTATCATCTCTGGAAAGAACCCCTACCCAGTCATTTATATACCCCACTGGG	7702
OY	5404	TTAAAGGCGCGGAGTGGAGCTTGATATCATCATCTTGAGTCAAAATCCGTGATACTGC	5463
Db	7703	TTAAAGGCGCGGAGTGGAGCTTGATATCATCATCTTGAGTCAAAATCCGTGATACTGC	7762
OY	5464	CAAGGAACCTGAAAAATCCAGAGACAACGCTAGCTATTCCTGTGAACCTCTAAGGATT	5523
Db	7763	CAAGGAACCTGAAAAATCCAGAGACAACGCTAGCTATTCCTGTGAACCTCTAAGGATT	7822
OY	5524	TGCGCCTGCTCTTAAACAAACACAGAGAGAAATGTAATCAATTAATCCCCCATG	5583
Db	7823	TGCGCCTGCTCTTAAACAAACACAGAGAGAAATGTAATCAATTAATCCCCCATG	7883
OY	5584	GSCTCCCTTATCATATTTTCTCTKATAGTSTTGTATACCTSTTTTCACTCTCATCTCA	5643
Db	7882	GSCTCCCTTATCATATTTTCTCTKATAGTSTTGTATACCTSTTTTCACTCTCATCTCA	7941
OY	5644	CCCCCTCCATGCGCGCTGATGACCAAGTAGTCCCTCTAACMAAGATTCTATGGAATG	5703
Db	7942	CCCCCTCCATGCGCGCTGATGACCAAGTAGTCCCTCTAACMAAGATTCTATGGAATG	8001
OY	5704	CAGCGTCCCGAAATATTGATGCCCATTCGTATAGAGTCTTSTTAAGGGAACCCCAAC	5765
Db	8002	CAGCGTCCCGAAATATTGATGCCCATTCGTATAGAGTCTTSTTAAGGGAACCCCAAC	8061
OY	5764	TTCACTGCGCCACACCCATATGCCCCCGAATCGTATAGAGTCTTSTTAAGGGAACCCCAAC	5823
Db	8062	TTCACTGCGCCACACCCATATGCCCCCGAATCGTATAGAGTCTTSTTAAGGGAACCCCAAC	8122
OY	5824	GCAAAATCTCATTTATGGAACAGAAAAATGATTAATCTAGTGTCTGTAGAGACTTGG	5883
Db	8122	GCAAAATCTCATTTATGGAACAGAAAAATGATTAATCTAGTGTCTGTAGAGACTTGG	8181
OY	5884	GTCACTGTCTGTGGACTTATCTTCAACCAATCTGTAATGTCGATGGGGGTGGAGTTCAA	5944
Db	8182	GTCACTGTCTGTGGACTTATCTTCAACCAATCTGTAATGTCGATGGGGGTGGAGTTCAA	8244
OY	5944	GATCAGGCAAGAAAAACATGTAAAAAGATTAATCTCCCACTCACCCGGGTATCATGGC	6003
Db	8242	GATCAGGCAAGAAAAACATGTAAAAAGATTAATCTCCCACTCACCCGGGTATCATGGC	8301
OY	6004	ACCTTAGCCCCCTTCAAGAGACTAGATCTTCAAAATCTATGAAAACCTTCCTGATCCAT	6065
Db	8302	ACCTTAGCCCCCTTCAAGAGACTAGATCTTCAAAATCTATGAAAACCTTCCTGATCCAT	8361
OY	6064	ACTGCGCTGTGAAGCTTATTTAATACCAACCTCACTGGGCTCCAGAGGTCTCGGCCCA	6122
Db	8362	ACTGCGCTGTGAAGCTTATTTAATACCAACCTCACTGGGCTCCAGAGGTCTCGGCCCA	8422
OY	6124	AAACCTATATCTGTGTGATAGTCCCTCCCTGAACTTCAGGCGATATGTTCAATCCCT	6183
Db	8422	AAACCTATATCTGTGTGATAGTCCCTCCCTGAACTTCAGGCGATATGTTCAATCCCT	8483
OY	6184	GTACCTGAACATGGAACAACTTCAGCACAGAAATTAACACCACTTCGGTTTATAGGA	6244
Db	8482	GTACCTGAACATGGAACAACTTCAGCACAGAAATTAACACCACTTCGGTTTATAGGA	8544
OY	6244	CCTCTGTTTCAATSTGGAAATTAACCATATCTCAACCTCACTGTGTAAATTTTATGC	6303
Db	8542	CCTCTGTTTCAATSTGGAAATTAACCATATCTCAACCTCACTGTGTAAATTTTATGC	8601
OY	6304	AATATTAATATCAACAACCACTCCCAATGCAATCAGTGGGTAACTCTCTCCACACAATA	6365
Db	8602	AATATTAATATCAACAACCACTCTCCCAATGCAATCAGTGGGTAACTCTCTCCACACAATA	8665
OY	6364	GTCTGCTTACCTTCAGGAATATTTTGTGTGTGTACTTACGCTATCGTTGTTGAAT	6422


```
Db      8662 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCGTTGTGAT 8721
Qy      6424 GGCCTTCAGAAATCATGTGCTTCCTCATTTCTTAGTGCCCGGATGRCATCAACT 6483
Db      8722 GGGCTTCAGAAATCATGTGCTTCCTCATTTCTTAGTGCCCGGATGRCATCAACT 8781
Qy      6484 GAACAAATTTATACAGTTATGTATCTAAGCCCGCAAAAGAGTACCCATCTT 6543
Db      8782 GAACAAATTTATACAGTTATGTATCTAAGCCCGCAAAAGAGTACCCATCTT 8841
Qy      6544 CTTTGTATAGAGAGAGAGTGTAGTCACTAGTACTGCACTTGGCGGTATACA 6603
Db      8842 CTTTGTATAGAGAGAGAGTGTAGTCACTAGTACTGCACTTGGCGGTATACA 8901
Qy      6604 ACCTCTACTAGTCTCTACTACAACTATCTCAAGAACTAAATGGGGAATGGAAAGGCGT 6663
Db      8902 ACCTCTACTAGTCTCTACTACAACTATCTCAAGAACTAAATGGGGAATGGAAAGGCGT 8961
Qy      6664 GCGGACTCCCTGTGCTACCTTGCAAGATCACTTAACCTCCAGCAGTATGCTTCTA 6723
Db      8962 GCGGACTCCCTGTGCTACCTTGCAAGATCACTTAACCTCCAGCAGTATGCTTCTA 9021
Qy      6724 AATCGAAGCTTTAGCTGTCTACCGCTGAGAGAGGGGAACTGTATTTTATGGG 6783
Db      9022 AATCGAAGCTTTAGCTGTCTACCGCTGAGAGAGGGGAACTGTATTTTATGGG 9081
Qy      6784 GAAGAATGCTTATTTATGTATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAAT 6843
Db      9082 GAAGAATGCTTATTTATGTATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAAT 9141
Qy      6844 CSAGATCGAATACAGTACAGAGAGAGAGAGCTTGAAACCTGAGCCCTGGGCTCTC 6903
Db      9142 CSAGATCGAATACAGTACAGAGAGAGAGAGCTTGAAACCTGAGCCCTGGGCTCTC 9201
Qy      6904 ACCGATGATGCGCTGAGATCTCCCTTTTGAAGCTTACAGCTATATATTTGCTA 6963
Db      9202 ACCGATGATGCGCTGAGATCTCCCTTTTGAAGCTTACAGCTATATATTTGCTA 9261
Qy      6964 CTCCTCTTGGAGCCCTGTATCTTTTACCTCCTGTTAACTTTTCTCTTCCAGAACTGA 7023
Db      9262 CTCCTCTTGGAGCCCTGTATCTTTTACCTCCTGTTAACTTTTCTCTTCCAGAACTGA 9321
Qy      7024 GCTGTAAATCTAACAAATGAGAGCCCAAGATGAGCTCCAGACTAAGCTTACCGCAGACC 7083
Db      9322 GCTGTAAATCTAACAAATGAGAGCCCAAGATGAGCTCCAGACTAAGCTTACCGCAGACC 9381
Qy      7084 CTGAGACCGGCTGTATGAGCCAGATCTGATGTTAATGACATCAAGAGCAACCTCTCTAG 7143
Db      9382 CTGAGACCGGCTGTATGAGCCAGATCTGATGTTAATGACATCAAGAGCAACCTCTCTAG 9441
Qy      7144 GAATCTCAGTGCACAACTCTACTAGAGCCCAATTCAGAGGAAAGAGTATGAGCGGT 7203
Db      9442 GAATCTCAGTGCACAACTCTACTAGAGCCCAATTCAGAGGAAAGAGTATGAGCGGT 9501
Qy      7204 SGTGCGCAACCTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGAATGAGAGAC 7263
Db      9502 C-TGCGCAACCTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGAATGAGAGAC 9560
Qy      7264 AGGATCTAGTATTTCTTGAAGCTATTAAGAACTCTTAAGCTTASGTGGAAAGTAC 7323
Db      9561 AGGATCTAGTATTTCTTGAAGCTATTAAGAACTCTTAAGCTTASGTGGAAAGTAC 9620
Qy      7324 ACATCACCCTTTAAACAGGAGGCTTGCAACTAGTCAACCTGACCAATCAGAGAGCT 7383
Db      9621 ACATCACCCTTTAAACAGGAGGCTTGCAACTAGTCAACCTGACCAATCAGAGAGCT 9680
Qy      7384 ACTATAATGCTTATAGGCAAGAGAGAGTAAAGAAATAGCAATCATATTTATGCTG 7443
Db      9681 ACTATAATGCTTATAGGCAAGAGAGAGTAAAGAAATAGCAATCATATTTATGCTG 9740
Qy      7444 AGAGCAGAGAGAGAGAGAGTATGAGTAAACCAAGTCTTTCAGAGCGGCAAGG 7503
Db      9741 AGAGCAGAGAGAGAGAGAGTATGAGTAAACCAAGTCTTTCAGAGCGGCAAGG 9800
```

```
Qy      7504 CAACCCCTTTGGGATCCCTCCCTTTGTATGAGGAGCTCTGTTTATGCTATTTCACTCT 7563
Db      9801 CAACCCCTTTGGGATCCCTCCCTTTGTATGAGGAGCTCTGTTTATGCTATTTCACTCT 9860
Qy      7564 ATTAATCTTGACCTGCR 7582
Db      9861 ATTAATCTTGACCTGCA 9879

RESULT 13
AX007980      10499 bp      DNA      linear      PAT 06-SEP-2000
LOCUS      AX007980
DEFINITION      Sequence 3 from Patent WO967395.
ACCESSION      AX007980
VERSION      AX007980.1 GI:9995677
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Perin,J.P., Rieger,F. and Alliel,P.M.
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs and their uses
Patent: WO 967395-A 5 29-DEC-1999.
JOURNAL      INST NAT SANTE RECH MED (FR); PERIN,JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
location/Qualifiers
FEATURES
source
1..10499
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="caxon:9606"

ORIGIN
Query Match      84.9%; Score 6436; DB 6; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

Qy      606 GATGGGAAACGTTCCCGCAAGACAAAACGCCCTAAGACGTATCTGABAAATTGGGA 665
Db      2953 GATGGGAAACGTTCCCGCAAGACAAAACGCCCTAAGACGTATCTGABAAATTGGGA 3012
Qy      666 MCAATTTGACCTCAGACATTAAGAAAGAAAGCACTTATTTCTTCTGACGTGCGGCTG 725
Db      3013 CCAATTTGACCTCAGACATTAAGAAAGAAAGCACTTATTTCTTCTGACGTGCGGCTG 3072
Qy      726 GCACTCCTGAGGAAAGTAAATTAATTAACACATCTTACAGCTAGACTTTTGTAGAA 785
Db      3073 GCACTCCTGAGGAAAGTAAATTAATTAACACATCTTACAGCTAGACTTTTGTAGAA 3132
Qy      786 AAGCAAAATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 845
Db      3133 AAGCAAAATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3192
Qy      846 AATTATGTAATAAATGTATTTATGCTTACAGAGAGCTTCAAGGTCTACTCCCTATC 905
Db      3193 AATTATGTAATAAATGTATTTATGCTTACAGAGAGCTTCAAGGTCTACTCCCTATC 3252
Qy      906 CCAGCATCCCGAGCTCTCTCCCAATTAATTAAGAGCCCTTCAACCAATGCTCA 965
Db      3253 CCAGCATCCCGAGCTCTCTCCCAATTAATTAAGAGCCCTTCAACCAATGCTCA 3312
Qy      966 AAGAGATAGCAAAAGGTTAAACAGTGAACCAAGAGTGCATATTTCCCAATTATGA 1025
Db      3313 AAGAGATAGCAAAAGGTTAAACAGTGAACCAAGAGTGCATATTTCCCAATTATGA 3372
Qy      1026 CCCCTCCCAAGCACTGGGAGAGAGATTCGAGCCGAGAGGAGAGAGTGTCTTTTXX 1085
Db      3373 CCCCT-CCAGCACTGGGAGAGAGATTCGAGCCGAGAGGAGAGTGTGTGCTTTTTC 3431
Qy      1086 TCTCCAGACTTAAAGCAATTAATAACAGACTTAGTAAATTCAGATAATCCTGATGG 1145
```

Db 3432 TCTCCAGACTTAAAGCAATTAABAAACAGACTTAAGTAATTCACAGATAACCCCTGATG 3491
Qy 1146 CTATATGTTGTTTTTACAGGGTTAGACAATCTTTGATCTGACATGAGAGATATATA 1205
Db 3492 CTATATGATGTTTTTACAGGGTTAGACAATCTTTGATCTGACATGAGAGATATATA -A 3550
Qy 1206 TGTCACTGCTAATTCAGACCTAACCCCAATAGAGAGAGTGCACCATATGCGAGCCT 1265
Db 3551 TGTCACTGCTAATTCAGACCTAACCCCAATAGAGAGTGCACCATATGCGAGCCT 3610
Qy 1266 GAGGTGTTGGGATCTCTGATATCTCACTGAGGTCAATGATNGATGACAAACAGAGG 1325
Db 3611 GAGAGTTGGGATCTCTGATATCTCACTGAGGTCAATG--ATAGATGACAAACAG-AGG 3667
Qy 1326 AAAAGANAATGATTTCCCAAGGCCAGACAGCAAGTTCACGCTTACACCTTCATTTGGGAGC 1385
Db 3668 AAGAGATGATTTCCCAAGGCCAGACAGCAAGTTCACGCTTACACCTTCATTTGGG- 3724
Qy 1386 ACAGAAATCAATACATGAGAGATGCTGACAGATTTGCTTAATTTGTTGCTASAA 1445
Db 3725 ACACAGATTCAGAACAT-GGAGATTTGCTGACAGATTTGCTTAATTTGTTGCTAGAA 3783
Qy 1446 GGACTAAGGAAAACTASAGAAABTCTAYGAATTACTCAATGATGTCACCAATACACA 1505
Db 3784 GGACTAAGGAAAACTAGAAAG-AGGTCTAGATTACTCAATGATGTCACCAATACACA 3842
Qy 1506 GGGGAAGGAAAAATCTACTGCTTTCTGAGAGACTAAGGAGGCAATTGAGAGC 1565
Db 3843 -GGGAAGGAAAAATCTACTGCTTTCTGAGAGACTAAGGAGGCAATTGAGAGC 3901
Qy 1566 GTGCTCTCTGTCACTGACTCTTCTGAGGCGCACTAATCTTAACCGTAAGTTTCA 1625
Db 3902 GTGCTCTCTGTCACTGACTCTTCTGAGGCGCACTAATCTTAACCGTAAGTTTCA 3961
Qy 1626 CTGAGTCACTGACAGATTAAG-AAAAAATTCAAAAGTCTGCGTACGCGCGGAGGAGAA 1684
Db 3962 CTGAGTCACTGACAGATTAAG-AAAAAATTCAAAAGTCTGCGTACGCGCGGAGGAGAA 4021
Qy 1685 ACTTAGAAACCTAATTAATGCGCAACTGCGTTTATTAATAGATCAGAGAGC 1744
Db 4022 ACTTAGAAACCTAATTAATGCGCAACTGCGTTTATTAATAGATCAGAGAGC 4081
Qy 1745 AGCGGGAACAGACAAACGGGATTAABAAAAGCGCCGCTTAAGTACAGACCTCAGG 1804
Db 4082 AGCGGGAACAGACAAACGGGATTAABAAAAGCGCCGCTTAAGTACAGACCTCAGG 4141
Qy 1805 CAAGTGAATTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTGATCCTAATAGG 1864
Db 4142 CAAGTGAATTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTGATCCTAATAGG 4201
Qy 1865 GCTTGCTTCAAGTGGGCTTCAAGAGCACTTTAAAAAAGTTGTCAGAGTAAG 1924
Db 4202 GCTTGCTTCAAGTGGGCTTCAAGAGCACTTTAAAAAAGTTGTCAGAGTAAG 4261
Qy 1925 CCGGCTTTGTCATGCGCTTTATTTCAAGGGAATCACTGGAAGGCCCATGCGCCAGG 1984
Db 4262 CCGGCTTTGTCATGCGCTTTATTTCAAGGGAATCACTGGAAGGCCCATGCGCCAGG 4321
Qy 1985 GGAAGAGTCTTTTGAAGTGAAGCACTAACAGATGATCCAGAGCAAGATGAGG 2044
Db 4322 GGAAGAGTCTTTTGAAGTGAAGCACTAACAGATGATCCAGAGCAAGATGAGG 4381
Qy 2045 TGCTTGAGGCAAGCGCATCCCATGCACTACCTCAGAGAGCCCTGGGTATGCTTGACC 2104
Db 4382 TGCTTGAGGCAAGCGCATCCCATGCACTACCTCAGAGAGCCCTGGGTATGCTTGACC 4441
Qy 2105 ATTGAAGGCAAGAGTGTCTCTGAGCACTGGTGGGCTTCTTAAGTCTAATCTC 2164
Db 4442 ATTGAAGGCAAG-AGGTGTCTCTGAGCACTGGTGGGCTTCTTAAGTCTAATCTC 4500
Qy 2165 TGTCCCGGCAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGCTCMTAAGCGGCA 2224
Db 4501 TGTCCCGGCAACTGTCTCTCAGATCTGTCACTA-TCGAGGGGGTCC-TAAGAGGGCA 4558

Qy 2225 GTCACTAGATCTTTTCTCCAGCACTAAGTTATGAAGTGGGAGCTTATCTTTTAC 2284
Db 4559 GTCACTAGATCTTTTCTCCAGCACTAAGTTATG-ACGGGAGGCTTATCTTTTAC 4616
Qy 2285 ATGCTTTTCAATTAATGCTTGAAGGCCCACTACCTTGTAGGAGACATTTCTAGCAA 2344
Db 4617 ATGCTTTTCAATTAATGCTTGAAGGCCCACTACCTTGTAGGAGACATTTCTAGCAA 4676
Qy 2345 AAGCAGGGGCAATTAATCACTGAAATGAGAGAGAAACCCCTTTGTTGTTGCCCTG 2404
Db 4677 AAGCAGGGGCAATTAATCACTGAAATGAGAGAGAAACCCCTTTGTTGTTGCCCTG 4735
Qy 2405 CTGAGAGAGAAATTAATCTGAAGCTGGGCAACAGAGACATATGACAGCCAA 2464
Db 4736 CTGAGAGAGAAATTAATCTGAAGCTGGGCAACAGAGACATATGACAG-CAAA 4794
Qy 2465 GAATGCCGCTCTGTTCAAGTTAACTAAGATTCACCTTCCCTACCAAGGCA 2524
Db 4795 GAATGCCGCTCTGTTCAAGTTAACTAAGATTCACCTTCCCTACCAAGGCA 4854
Qy 2525 GTAACCCCTCAAGCCCAAGGCCCAACAGATTCAAAAGATTTGAAGACTTAAAGC 2584
Db 4855 GTAACCCCTCAAGCCCAAGGCCCAACAGATTCAAAAGATTTGAAGACTTAAAGC 4914
Qy 2585 CCAAGGCTAATTAACCAATGACATACTCCCTGACATTAATCCGTAGTGAGGAGG 2644
Db 4915 CCAAGGCTAATTAACCAATGACATACTCCCTGACATTAATCCGTAGTGAGGAGG 4967
Qy 2645 CACAGAAACCAAGTGAACAGTGGGTTAAGTCAAGATTCAGATTAATCAATGAGGC 2704
Db 4968 TACAGAAACCAACAGAGTGA-AGTTAGTCAAGATTCAGATTAATCAAT-AGGC 5025
Qy 2705 CGTTGCTTTTATACCAAGCTGTAACCTTATCTGCTGCTGCTTCCAAATACAGA 2764
Db 5026 TGTGTTTCTCTATACCAAGCTGTAACCTTATCTGCTGCTTCCAAATACAGA 5085
Qy 2765 GGAAGAGAGTGGTTTAACTGACCTGGAACCTTMAAGATGCTTCTGCAATCCCTGACA 2824
Db 5086 GGAAGAGAGTGGTTTAACTGACCTGGAACCTTMAAGATGCTTCTGCAATCCCTGACA 5145
Qy 2825 TCTGACTCTCAATTTGTTTGGCTTGAAGTACTTCAACCCATCTCACTC 2884
Db 5146 TCTGACTCTCAATTTGTTTGGCTTGAAGTACTTCAACCCATCTCACTC 5205
Qy 2885 CTGACTTTTACCCCAAGGTTCAAGGATTAATCCCATCTAATTTGGCCAGGCAATAGC 2944
Db 5206 CTGACTTTTACCCCAAGGTTCAAGGATTAATCCCATCTAATTTGGCCAGGCAATAGC 5265
Qy 2945 CCAAGACTGAGTCATTTTCAATACCTGACACTGTTGCTCTCTGTAAGTGAATTT 3004
Db 5266 CCAAGACTGAGTCATTTTCAATACCTGACACTGTTGCTCTCTGTAAGTGAATTT 5323
Qy 3005 ACTTTTGGCCGCTTCAAGAACTTTGCTCATCAAGCAACCCAGGCTCTTAAATTT 3064
Db 5324 ACTTTTGGCCGCTTCAAGAACTTTGCTCATCAAGCAACCCAGGCTCTTCAATTT 5383
Qy 3065 CTTGCTAATCTGTTGCTACAGGTTTCCAAACGABAGTCTCAPCTGCTCAGAGAGG 3124
Db 5384 CTTGCTAATCTGTTGCTACAGGTTTCCAAACGABAGTCTCAPCTGCTCAGAGAGG 5443
Qy 3125 TAAATTAATGATTAATTAATCAAGGCAACAGGCTCTGAGTGAAGTAATCA 3184
Db 5444 ----TACTTGAAGGCTAATTAATTAATCAAGGCAACAGGCTCTGAGTGAAGTAATCA 5499
Qy 3185 GCTAATATGCTTATCTCATCTCAATCAACCTTAAGCAACTAAGGRTTCTTTGGCT 3244
Db 5500 GCTAATATGCTTATCTCATCTCAATCAACCTTAAGCAACTAAGGRTTCTTTGGCT 5559
Qy 3245 AATAGGTTTCTGCGAAGATGATTCCTCAGGCTTGGCTAATGCAAGGCTATTAATA 3304
Db 5560 AATAGGTTTCTGCGAAGATGATTCCTCAGGCTTGGCTAATGCAAGGCTATTAATA 5618

Oy	3305	CASATTAATGAGAACTCGAAGAACCCAAATACCACTTTATTATAGTGAAYAMCTGAAGYMR	3364
Db	5619	CACATAATTAAGGAAACTCAGAAAAGCCAAATCCCATTTAGTAAGATGACAACCTGAAGAATG	5678
Oy	3365	AAGTGCTTTCCAGAGGCCCTTAAAGAAAGGCCTTAAACCCAAGYCCCACTGTTAATYTGGC	3424
Db	5679	AAGTGGCTTTCCAAGGCC-----TAACCCAAAGCCCCAGTGTTAAGTTGGC	5724
Oy	3425	AAACRGGCAGACTTTTSTTYATATATTCACAGAAAAAAACAGRAAYAGCTCTRGAGATCC	3484
Db	5725	AAACAGGGCAGACTTTCTTCATATGTACAG-AAAAACAGGAATAGCTCTGAGATCC	5783
Oy	3485	TTCACACATCCRAGGAGAGAGCTTGCCAACTVTGGCRYACTGTASTAGGAAAYTGATG	3544
Db	5784	TTACACAAATCCGAGGAGTAGCTTGCCAACTGTGGCATCTCGACTTAAGAAATTGATG	5843
Oy	3545	TAGTGGCAAAAGGCTGRCYTCAATGTGTTATYGGGTAGGGTGGGAGTAGAGAGTAKTAT	3604
Db	5844	TAGTGGCAAAAGGTTGACTCTCATTTGTTTACGGGTAGTGTGGCAGTAGAGCTGTAT	5903
Oy	3605	CTGAAGCAGTTTAAATTAATACAGGGRAGAGATCTTACTGTGTGACATCTCATGAGTGA	3664
Db	5904	CTGAAGCAGTTTAAATTAATACAGGGAGAGATCTTACTGTGTGACATCTCATGAGTGA	5963
Oy	3665	AYRCATATCTCACTGTCTAAAGAGACTTGTGGCTGTCAAGCAACVETTTTACTTAAATRTC	3724
Db	5964	ATGGCATACTACCTAGCTTAAAGAGACTGTGTGCTGTCAAGCAACTGTTTTACTTAAATRTGTC	6023
Oy	3725	AGGCTCATTACTTGAARGGCCAGAGCTGCRACACTGGCACTGTGCACTCTTAAACCCAG	3784
Db	6024	AGGCTCATTTACTTGAAAGGGCCAGTGTGCGACTGTGCACTGTGTGCACTCTTAAACCCAG	6083
Oy	3785	YCNCATTTCTTCCAGACAATGAGAAAGATRAAYATTACTGTCAACARTATTTCTC	3844
Db	6084	CCACATTTCTTCCAGACAATGAGAAAGATRAAATTAACGTCAACAAGTATTTCTC	6143
Oy	3845	AAACCTATNGCACCTCGAGGGACCTTYYNAGAGTTCTCTTACTGATCCYAGCTTCCAA	3904
Db	6144	AAACCTATNGCACCTCGAGGGACCTTYYNAGAGTTCTCTTACTGATCCCGAAC-TCAAC	6202
Oy	3905	TTGTATCTGATGAGAGTTCCCTTGTAGAAAAAGACTTGC AAAAGYGSGGNATGCAAG	3964
Db	6203	TTGTATCTGATGAGAGTTCTCTTGTAGAAAAAGACTTGC AAAAGYGSGGNATGCAAG	6262
Oy	3965	GTCAGTATATGGAATATTTGAAGTATCCCTCACTCCAGGAACCTAGTCTYATGCR	4024
Db	6263	GTCAGTATATGGAATATCTTGAAGTATCCCTCACTCCAGGAACCTAGTCTCACTGCR	6322
Oy	4025	GCAGAACTAATAGCCCTCAATYKGGGCACCTAGATTTAGAGAGAGRAAAAAGGYAATATAT	4084
Db	6323	GCAGAACTAATAGCCCTCACTYGGGCACCTAGATTTAGAGAGAGRAAAAAGGCCAATATAT	6382
Oy	4085	TATACAGACTCTBARTATAGCTYACCTTGMTCATGCCATGCMGCAATATGSRAPAGAA	4144
Db	6383	TATACAGACTCTAATATATAGCTTACTGTCTCTCCATGCCATGCAATATAGGAAAGAA	6442
Oy	4145	AGGGAATTCCTAATCTTCYAGRGAACCTATCAMACATCGAGGAAGCACTTAGABATTA	4204
Db	6443	AGGGAATTCCTAATCTTCGAGAGAACCTATCAMACATCGAGGAAGCACTTAGGAAATTA	6502
Oy	4205	TTATYTGCGWTPACAGAAACCTAPADAGGTGMACTTTACACTGCGYGGGTCATCANAAA	4264
Db	6503	TTATYTGCTGTACAGAAACCTAAPAGAGTGGCAGCTTACACTGCGGGGTCTACAGAAA	6562
Oy	4265	GGAAARBAAGGGAAATASAAGRBAATTCGCCAAGCARATATTGAAGCMAAAAAGACTGCA	4324
Db	6563	GGAAARBAAGGGAAATRAGAGAACTGCCAACAGTATTTGAAGCCAAAAAGACTGCA	6622
Oy	4325	AGCAGAGACCTTCATTAGAAATGCTTATTTAACTTCCCTTAGTATAGGTAATCCCTTC	4384
Db	6623	AGCAGAGACCTTCATTAGAAATGCTTATTTAAACAACCTTCATGTRTAGGTAATCCCTTC	6682
Oy	4385	CGGGAACCAAGCCCCAGTACTACGACGAGAGAAAACAGATGGGGAACCTCAAGAG-CAG	4443

Db	6683	CGGGAAACCAAGCCCCAGATCTAGAGAGGAAACGAATGGGGAACTTCACAGAGACAG	674
Qy	4444	TTTTCTCCCTCGGAGCGGTTTACCTGACCTGAAGAGGAGAAATACCTTTTGGCTGCACTAT	450
Db	6743	TTTTCTCCCTCGGAGCGGCTGACCTGACCTGAGAGAGGAGAAATACCTTTTGGCTGCACTAT	680
Qy	4504	CCATGGAATTACTTAAACCTTCATCAAACTTTCACCTTAGGCGATGATGACCCA	456
Db	6803	CCATGGAATTACTTAAACCTTCATCAAACTTTCACCTTAGGCGATGATGACCCA	686
Qy	4564	TCARATGGCCAAATCATTATTTTACTGAGACCAAGGCTTTTCAAACTATCAAGCAATAT	462
Db	6863	TCGATATGGCCAAATCATTATTTTACTGAGACCAAGGCTTTTCAAACTATCAAGCAATAT	692
Qy	4624	CAGGCGCTGTGAATGTGCGACABAAATATATCCCGCTCTATATCGGCAAGCTCTTCAAG	468
Db	6923	CAGGCGCTGTGAATGTGCGACABAAATATATCCCGCTCTATATCGGCAAGCTCTTCAAG	698
Qy	4684	ARAAACAAPDAACGAGCCATTACCTCTGAPAPABACTGGCAACTGATTTTACCAACAAGCC	474
Db	6983	AGAAACAAPDAACGAGCCATTACCTCTGAPAPABACTGGCAACTGATTTTACCAACAAGCC	704
Qy	4744	AAACCTCAGGGAATTCAGATCTTACTAGCTGGGTABATACCTTTCACGGGTTGGCAGAG	480
Db	7043	AAACCTCAGGGAATTCAGATCTTACTAGCTGGGTABATACCTTTCACGGGTTGGCAGAG	710
Qy	4804	GCCTTCCCTGTAGAGACAGAAAGGCCCAAGAGTAAATPAAAGCACTAGTTCATGAAATA	486
Db	7103	GCCTTCCCTGTAGAGACAGAAAGGCCCAAGAGTAAATPAAAGCACTAGTTCATGAAATA	716
Qy	4864	ATTCGCCAGTTCGGACTTCCCGGAGCTTACAGAGTGAACAATAGCCCTGCTTTCAGAGCC	492
Db	7163	ATTCGCCAGTTCGGACTTCCCGGAGCTTACAGAGTGAACAATAGCCCTGCTTTCAGAGCC	722
Qy	4924	AACAGTAACCCAGGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGGCTGA	498
Db	7223	AACAGTAACCCAGGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGGCTGA	728
Qy	4984	AGGCCACAGTCTCTCAGGAGAGGTGACAGAAATGAAATGAAATGAAATGAAATGAAATGAA	504
Db	7283	AGGCCACAGTCTCTCAGGAGAGGTGACAGAAATGAAATGAAATGAAATGAAATGAAATGAA	734
Qy	5044	AAGCAACCCAGAGAAACCACTCAATGAGGCTGTCTGTTGCCATATAGCCTTAAAGAA	510
Db	7343	AAGCAACCCAGAGAAACCACTCAATGAGGCTGTCTGTTGCCATATAGCCTTAAAGAA	740
Qy	5104	ATTCGCACTTTCCCAAAAAGCAGAGCTTAGCCATACGAAATGCTGTATGAGAAAGCCC	516
Db	7403	ATTCGCACTTTCCCAAAAAGCAGAGCTTAGCCATACGAAATGCTGTATGAGAAAGCCC	746
Qy	5164	TTCATPACCAATGACCTTGCTGATGACCAAGACAGGCAACTTATGTTGACACATCACT	522
Db	7463	TTCATPACCAATGACCTTGCTGATGACCAAGACAGGCAACTTATGTTGACACATCACT	752
Qy	5224	CCTTAGCCAAATATCAACAGTCTTAAACATTAACAAGAACCTATCCCTGAGAAAGAG	528
Db	7523	CCTTAGCCAAATATCAACAGTCTTAAACATTAACAAGAACCTATCCCTGAGAAAGAG	758
Qy	5284	GAAAAAGAACTTTCACCCGAGGACATGATTTAGTCAAGTCCCTTCTCTAATTC	534
Db	7583	GAAAAAGAACTTTCACCCGAGGACATGATTTAGTCAAGTCCCTTCTCTAATTC	764
Qy	5344	CATCCCTAGATACATCCGAGGAGAGACCTTACCCAGTCACTTTATATTAATCCCAACTG	540
Db	7643	CATCCCTAGATACATCCGAGGAGAGACCTTACCCAGTCACTTTATATTAATCCCAACTG	770
Qy	5404	TTAAAGTGGCTGAGTGAAGTCTTGATATATCACTGATGATGATGATGATGATGATGATG	546
Db	7703	TTAAAGTGGCTGAGTGAAGTCTTGATATATCACTGATGATGATGATGATGATGATGATG	776
Qy	5464	CAAGGAACCTGAAATTCAGAGACAGAGCTGATTTCTGTTGAATCTTATGAGATTT	552

Db 7763 CAAAGAACTGAAATCCAGAGACAACGCTAGCTATTCCTGTGAAACCTCTAGAGATT 7822
Qy 5524 TGGCGCTGCTCTTCAAAACAACACAGAGGAAAAGTAATTAATAATCATTAATCCCCATG 5583
Db 7823 TGGCGCTGCTCTTCAAAACAACACAGAGGAAAAGTAATTAATAATCATTAAT-CCCCATG 7881
Qy 5584 GSGCTCCCTTATCATATTTTCTCTKTAATGTTSTTTTAAAGCTSTTTTCACTCTACATGCA 5643
Db 7882 GCGCTCCCTTATCATATTTTCTCTTATCTGTTCTTTTAAACCTCTTTCACCTCACTGCA 7941
Qy 5644 CCCCCCATGCGCTGTATGACCAAGTAGCTCCCTTACCAAGATTCTATGAGAAATG 5703
Db 7942 CCCCCCATGCGCTGTATGACCAAGTAGCTCCCTTACCAAGATTCTATGAGAAATG 8001
Qy 5704 CAGCGTCCCGGAAATATTGATGCCCATGTATAGAGAGTCTTTSTTAAGGAAACCCCAAC 5763
Db 8002 CAGCGTCCCGGAAATATTGATGCCCATGTATAGAGAGTCTTTSTTAAGGAAACCCCAAC 8061
Qy 5764 TTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTGCACTCTTTTGATGAT 5823
Db 8062 TTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTGCACTCTTTTGATGAT 8121
Qy 5824 GCAATATCTCATTTTGGACAGAGAAAATGATTAATCTAGTTGCTGAGAGACTTGA 5883
Db 8122 GCAATATCTCATTTTGGACAGAGAAAATGATTAATCTAGTTGCTGAGAGACTTGA 8181
Qy 5884 GTCACTGTGTGGACTTACTTCAACCAACCTGTATGTGTATGAGGGGTGAGTTCA 5943
Db 8182 GTCACTGTGTGGACTTACTTCAACCAACCTGTATGTGTATGAGGGGTGAGTTCA 8241
Qy 5944 GATCAGGCGAAGAAAACATGTAAAGAAATCTCCCACTCACGSGGTATCATAGC 6003
Db 8242 GATCAGGCGAAGAAAACATGTAAAGAAATCTCCCACTCACGSGGTATCATAGC 8301
Qy 6004 ACCTCTAGCCCTCTCAAAAGACAGATCTCAAAACTATGAAACCTCCGTACCCAT 6063
Db 8302 ACCTCTAGCCCTCTCAAAAGACAGATCTCTCAAAACTATGAAACCTCCGTACCCAT 8361
Qy 6064 ACTGCGCTGTGAAGCTTATTAATACACCTCTCACTGGGCTCATGAGAGTCTCGGCCAA 6123
Db 8362 ACTGCGCTGTGAAGCTTATTAATACACCTCTCACTGGGCTCATGAGAGTCTCGGCCAA 8421
Qy 6124 AACCTTACTAGTGTGATATGCTCCCTGAACTTCAAGGCAATATGTTCAATCCCT 6183
Db 8422 AACCTTACTAGTGTGATATGCTCCCTGAACTTCAAGGCAATATGTTCAATCCCT 8481
Qy 6184 GTACCTGAAACATGGAACAACTTACGACAGAAATTAACCACTTCCGTTTATGAGGA 6243
Db 8482 GTACCTGAAACATGGAACAACTTACGACAGAAATTAACCACTTCCGTTTATGAGGA 8541
Qy 6244 CTTCTTGTTCATATSTGAAATTAACCACTTCACTCACTCTGTGTAAATTTAGC 6303
Db 8542 CTTCTTGTTCATATSTGAAATTAACCACTTCACTCACTCTGTGTAAATTTAGC 8601
Qy 6304 AATATCATATACAAACCACTCCCAATGCACTAGGSGTAACCTCCCAACAATA 6363
Db 8602 AATATCATATACAAACCACTCCCAATGCACTAGGSGTAACCTCCCAACAATA 8661
Qy 6364 GTGCGCTACCTCAGAAATATTTTGTCTGTGTAACCTCAGCCTATGTTTGAAT 6423
Db 8662 GTGCGCTACCTCAGAAATATTTTGTCTGTGTAACCTCAGCCTATGTTTGAAT 8721
Qy 6424 GSGCTTTCAGAAATATATGCTCTCTCATTTCTTAATGCCCCATATGCACT 6483
Db 8722 GSGCTTTCAGAAATATATGCTCTCTCATTTCTTAATGCCCCATATGCACT 8781
Qy 6484 GAAACAATTTTATACATTAATGCAATATAGCCCCGCAACAAAAGATACCACTTCT 6543
Db 8782 GAAACAATTTTATACATTAATGCAATATAGCCCCGCAACAAAAGATACCACTTCT 8841
Qy 6544 CTTTGTGTAATAGAGACAGAGAGTGTAGTGAATGCACTAGTACTGCAATTTGGCGTATCA 6603
Db 8842 CTTTGTGTAATAGAGACAGAGAGTGTAGTGAATGCACTAGTACTGCAATTTGGCGTATCA 8901

Qy 6604 ACCTTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGAAACGGGTC 6663
Db 8902 ACCTTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGAAACGGGTC 8961
Qy 6664 GCGGACTCCCTGTGTCACTTGCAGATCACTTAATCTCCCTAGACAGATAGTCTTGRA 6723
Db 8962 GCGGACTCCCTGTGTCACTTGCAGATCACTTAATCTCCCTAGACAGATAGTCTTCAA 9021
Qy 6724 AATCGAAGACTTAACTACTGTCAACCGCTGAAAGAGGGGGAACCTGTTATTTTGAAG 6783
Db 9022 AATCGAAGACTTAACTACTGTCAACCGCTGAAAGAGGGGGAACCTGTTATTTTGAAG 9081
Qy 6784 GAAAGATGCTGTTATTAATGTTAATCAATCCGAATCGTCACTGGAAGATTAAAGAAAT 6843
Db 9082 GAAAGATGCTGTTATTAATGTTAATCAATCCGAATCGTCACTGGAAGATTAAAGAAAT 9141
Qy 6844 CSAGATCGAATACAGCTAAGAGAGAGAGACTTGAACACTGGAACCTGGGGCTCTC 6903
Db 9142 CSAGATCGAATACAGCTAAGAGAGAGAGACTTGAACACTGGAACCTGGGGCTCTC 9201
Qy 6904 AGCGATATGATGCCCTGGAATTCCTCCCTTTTGGAGACTCTGAGCGTATTAATTTGCTA 6963
Db 9202 AGCGATATGATGCCCTGGAATTCCTCCCTTTTGGAGACTCTGAGCGTATTAATTTGCTA 9261
Qy 6964 CTCCTCTTGGAGCCCTGTATCTTTTACCTCCTTGTAACTTGTCTCTTCAGAAATCGAA 7023
Db 9262 CTCCTCTTGGAGCCCTGTATCTTTTACCTCCTTGTAACTTGTCTCTTCAGAAATCGAA 9321
Qy 7024 GCTGTAAACTTACAAATGAGAGAGAGAGATGAGTCTCAAGACTTAAGTCTACCGAGACC 7083
Db 9322 GCTGTAAACTTACAAATGAGAGAGAGAGATGAGTCTCAAGACTTAAGTCTACCGAGACC 9381
Qy 7084 CTGAGCGGCTGTGTACCCACAGATCTGATGTTAATGACATCAAGGCAACCTCTGAG 7143
Db 9382 CTGAGCGGCTGTGTACCCACAGATCTGATGTTAATGACATCAAGGCAACCTCTGAG 9441
Qy 7144 GAAATCTCAGCTGACAACTCTTACTACGCCCCAATTCAGCAGGAAGCATTAAGCGGT 7203
Db 9442 GAAATCTCAGCTGACAACTCTTACTACGCCCCAATTCAGCAGGAAGCATTAAGCGGT 9501
Qy 7204 SGTGCGCAACCTCCCAACAGCACTTAAGTTTCTGTTGATGAGTGGGGGCTGAGAGAC 7263
Db 9502 SGTGCGCAACCTCCCAACAGCACTTAAGTTTCTGTTGATGAGTGGGGGCTGAGAGAC 9560
Qy 7264 AGAGCTAGCTGATTTCTTGGGCTGATTAAGATCCCTTAAGCTAGTGGGAAAGGTGACC 7323
Db 9561 AGAGCTAGCTGATTTCTTGGGCTGATTAAGATCCCTTAAGCTAGTGGGAAAGGTGACC 9620
Qy 7324 ACATTCACCTTTAAACAGGGGCTTGGCACTTAAGTCAACCTGACCAATCAGAGAGCTC 7383
Db 9621 ACATTCACCTTTAAACAGGGGCTTGGCACTTAAGTCAACCTGACCAATCAGAGAGCTC 9680
Qy 7384 ACTAAATGCTTAATTAAGCAAGAGAGAGAGGTTAAAGAAATGACCAATCATTAATGCMTG 7443
Db 9681 ACTAAATGCTTAATTAAGCAAGAGAGAGAGGTTAAAGAAATGACCAATCATTAATGCMTG 9740
Qy 7444 AGAGCAAGCAAGAGAGAGCAATGATCCGGATATTAACCAAGTATTCAGCGGCAACGG 7503
Db 9741 AGAGCAAGCAAGAGAGAGCAATGATCCGGATATTAACCAAGTATTCAGCGGCAACGG 9800
Qy 7504 CAACCCCTTTGGGCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db 9801 CAACCCCTTTGGGCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 9860
Qy 7564 ATTAATCTTGCATCTGCR 7582
Db 9861 ATTAATCTTGCATCTGCA 9879

RESULT 14
AY101586
LOCUS

AY101586 10229 bp DNA linear PRI 11-FEB-2004

DEFINITION Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1
locus, allele A, complete sequence.
ACCESSION AY101586
VERSION AY101586.1 GI:37544407
KEYWORDS
SOURCE
ORGANISM Pan troglodytes (chimpanzee)
REFERENCE
AUTHORS Pan troglodytes
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 10229)
1 (bases 1 to 10229)
Mallet, F., Boucon, O., Prudhomme, S., Cheynet, V., Oriol, G.,
Bonnard, B., Lucotte, G., Duret, L. and Mandrand, B.
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
1457826
JOURNAL 2 (bases 1 to 10229)
PUBMED Mallet, F., Boucon, O. and Oriol, G.
REFERENCE Direct Submission
AUTHORS Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-Biometrieux, Ecole Normale Supérieure de Lyon - 46 allée
JOURNAL d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10229
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/isolate="1"
/isolation_source="PBMC of chimpanzee 1"
/db_xref="taxon:9598"
/cell_type="PBMC"
/note="Isolated by PCR, ERVW1 locus, allele A
endogenous virus: ERV-W"
1..779
join(248..810,2889..9820)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..810,7592..9820)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..810,9255..9820)
/note="putative mRNA transcript 3; no coding region"
7819..9435
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA068165.1"
/db_xref="GI:37544408"
/translation="MALPYHIFLFTVILBSFTLAPPPRCMTSSPFYQELRMORP
GNIDAPSTKSKGPTFTAHTHMPNCHSATLCMAHANTHTWTKIIMPSCPGILG
TVCWTFTHTGMSDGGVQDQARERKVEKVIISQLTRHSASRYPKGLDLSKLTART
HTRVLSFNTTLTGLHEVSAONPTNCMILPLNFRPVYSIPIVEQNNSTINTTSV
LVGLPVSNEIHTSNLTCVFESNTDTSOCIRHVPPTQIVCLPSGIPIVCGTSA
YRLNGSSSESMGFLSFLVPMPTIYBODLYNVVSKPRKRPILIPVIGAVLALG
TGIGITSTOFTYKLSOELNDMERVADSLVTLDDQLNSLAAYLQNRALDLTAE
RSGTFLRGECCITVNSGIVTEKSEIRDRIRKRAEIRNTGFMGLSOMWPLP
FUGPLAAILLLFGPCIFNLLVNEFVSEIRIEAVKIQMEPKMOSKTKIYRRPDRPAS
RSDVNDIGTPEEIRITQPLRPNSAGSS"
9494..10229
ORIGIN
LTR
Query Match 83.7%; Score 6346.4; DB 9; Length 10229;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 6625; Conservative 194; Mismatches 110; Indels 52; Gaps 28;
QY 606 GATGGGAACGTTCCCGGACAGACAAAGCGCCCTAAGAGCTATTCTGGAAATTTGGGA 665
DB 2888 GATGGGAACGTTCCCGGACAGACAAAGCGCCCTAAGAGCTATTCTGGAAATTTGGGA 2947
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGACTTATTTCTTCTGAGTGGCCGCTG 725
DB 2948 CCAATTTGACCTCAGACACTAAGAAAGAAAGACTTATTTCTTCTGAGTGGCCGCTG 3007
QY 726 GCATCTCTGAGGGAAGTAAATTTAAACACATCTTACAGTGAAGCTTTTGTAGA 785

DB 3008 GCATCTCTGAGGGAAGTAAATTTAAACACATCTTACAGTGAAGCTTTTGTAGA 3067
QY 786 AAGGCAATGAGTGAAGTCCCAATAGTACAACTTTCTTTTCAATTAAGACACTCAC 845
DB 3068 AAGGCAATGAGTGAAGTCCCAATAGTACAACTTTCTTTTCAATTAAGACACTCAC 3127
QY 846 AATTATGTAATAAGTGTGATTTATGCGCTACAGAAAGCTTCAAGAGTCTACCTCCATC 905
DB 3128 AATTATGTAATAAGTGTGATTTATGCGCTACAGAAAGCTTCAAGAGTCTACCTCCATC 3187
QY 906 CCAGCATCCCC--GACTCTTCCCAATTAATAGGACCCCTTCAACCAATGCTGCC 963
DB 3188 CCAGCATCCCCCTGACTCTTCCCAATTAATAGGACCCCTTCAACCAATGCTGCC 3247
QY 964 AAAAGGATGATGACAAAGGATTAACAGTGAACCAAGATGCAATATTTCCCAATAT 1023
DB 3248 AAAAGGATGATGACAAAGGATTAACAGTGAACCAAGATGCAATATTTCCCAATAT 3307
QY 1024 GACCCCTCCCAAGCAGTGGGAGGAGAAATTCGCGCCAGCCAGAGTGCATGCTTTT 1083
DB 3308 GACCCCT-CCAGCAGTGGGAGGAGAAATTCGCGCCAGCCAGAGTGCATGCTTTT 3366
QY 1084 YTTCTCCAGACTTAAAGCAATTAAGCAAGCTTAGTAAATTTCTAGATAATCTGAT 1143
DB 3367 TCTCTCTCAGACTTAAAGCAATTAAGCAAGCTTAGTAAATTTCTAGATAATCTGAT 3426
QY 1144 GGCCTATATGATGTTTAAAGGTTAGACAAATCTTTGATCTGACATGAGAGATTA 1203
DB 3427 GGCCTATATGATGTTTAAAGGTTAGACAAATCTTTGATCTGACATGAGAGATTA 3486
QY 1204 TATGTCACTGTAAATCAGACACTAACCCCAATGAGAGAGTGCACCATTAATCTGACG 1263
DB 3487 -ATGTCACTGTAAATCAGACACTAACCCCAATGAGAGAGTGCACCATTAATCTGACG 3545
QY 1264 CTGAGGTTTGGCATCTCTGTATCTCAGTCAAGTCAATGATANGATGACACAGAA 1323
DB 3546 CCAGAGTTTGGCATCTCTGTATCTCAGTCAAGTCAATGATANGATGACACAG-A 3602
QY 1324 GGAAGANATATTTCCCAAGGCGCAGACAGACATGTCCTCACTCACTCACTGGG 1383
DB 3603 GGAAGAGATGATTTCCCAAGGCGCAGACAGATGTCCTCACTCACTCACTGGG- 3661
QY 1384 ACACAGAAATCAGTAACATGGAAGATGTTGGTGTGAGACATTTGCTCAATCTGATGCTAS 1443
DB 3662 --ACACAGAAATCAGTAACATGGAAGATGTTGGTGTGAGACATTTGCTCAATCTGATGCTAS 3718
QY 1444 AAGACTAAGGAAAATCTASGAAAGAAATCTAAGAAATCTCAATGATGCTCAACATACA 1503
DB 3719 AAGACTAAGGAAAATCTASGAAAG--AAGCTAATGAAATTAATCAATGATGCTCAATACA 3777
QY 1504 CAGGGGAAAGGAAAGAAATCTCACTGCTTTCTGAGAGACTAAGGAGGCAATTGAGAA 1563
DB 3778 CA-GGAAAGGAAAGAAATCTCACTGCTTTCTGAGAGACTAAGGAGGCAATTGAGAA 3836
QY 1564 GCGTCCCTCTGCTGACCTGACTCTTGAAGGCGCAATCTTAAGCGTAAGTTAT 1623
DB 3837 GCGTCCCTCTGCTGACCTGACTCTTGAAGGCGCAATCTTAAGCGTAAGTTAT 3896
QY 1624 CACTCAGTCACTGACATTAAG-AAAAAATTTCAAAAGTCTGCGTAGGCGCGAGCA 1682
DB 3897 CACTCAGTCACTGACATTAAGAAAAAATTTCAAAAGTCTGCGTAGGCGCGAGCA 3956
QY 1683 AAACCTTGAAGAACCTTATGAACTTGGCAACTGCTTTTATATAAGATCAGAGAGA 1742
DB 3957 AAACCTTGAAGAACCTTATGAACTTGGCAACTGCTTTTATATAAGATCAGAGAGA 4016
QY 1743 GCAGCGGGAACAGAGCAACCGGATTAAGAAAAAGGCGCACCGCTTATGATGACCTCA 1802
DB 4017 GCAGCGGGAACAGAGCAACCGGATTAAGAAAAAGGCGCACCGCTTATGATGACCTCA 4076
QY 1803 GCGAAGTGAATTTGAGGCTCTGGAAGGGAAGGCTGGGCAATTAATGCTTAATA 1862
DB 4077 GCGAAGTGAATTTGAGGCTCTGGAAGGGAAGGCTGGGCAATTAATGCTTAATA 4136

OY 1863 GGGCTTGCTTCAGTGGGCTTACAAAGACATTTAAAAAGTTGTCCAGTAGAAGTA 1922
DB 4137 GGGCTTGCTTCAGTGGGCTTACAAAGACATTTAAAAAGTTGTCCAGTAGAAGTA 4196
OY 1923 AGCGCGCCCTTGCTTCAGTGGGCTTACAAAGGAAATCATCTGGAAGGCCCATGCGCCA 1982
DB 4197 AGCGCGCCCTTGCTTCAGTGGGCTTACAAAGGAAATCATCTGGAAGGCCCATGCGCCA 4256
OY 1983 GGGGCAAAAGGCTTTTGAAGTCAAGAACCACTAACAGATGATCCAGCAGCAGACTGAG 2042
DB 4257 GGGGCAAAAGGCTTGAAGTCAAGAACCACTAACAGATGATCCAGCAGCAGACTGAG 4316
OY 2043 GGTGCTGGGGCAAGGCGCATCCCATGCGCATACCTCTACAGAGCCCTGGGTAAGTTGA 2102
DB 4317 GGTGCTGGGGCAAGGCGCATCCCATGCGCATACCTCTACAGAGCCCTGGGTAAGTTGA 4376
OY 2103 CCATGAGGGGCGAGGAAGGTTGTCTCTGGAACATGAGGGGCTTCTTACTTACTCT 2162
DB 4377 CCATGAGGGGCGAGG-AGGTTGTCTCTGGAACATGAGGGGCTTCTTACTTACTCT 4435
OY 2163 TCTGTCCCGGCAACTGTCTCCAGATCTGTCACTATCTGAGGGGCTCCTTAAGAAGG 2222
DB 4436 TCTGTCCCGGCAACTGTCTCCAGATCTGTCACTATCTGAGGGGCTCCTTAAGAAGG 4493
OY 2223 CAGTCACTAGATATCTTTTCCAGCCACTAAGTATGAACTGGGGAGCTTTATCTTTTC 2282
DB 4494 CAGTCACTAGATATC-TTCTCCCGGCACTAAGTATG-ACTGGGGAGCTTTATCTTTTC 4551
OY 2283 ACATGCTTTCTAATATGCTTAAAGCCCACTACCTTGTGAGGAGAGACATTTAG 2342
DB 4552 ACATGCTTTCTAATATGCTTAAAGCCCACTACCTTGTGAGGAGAGACATTTAG 4611
OY 2343 AAAAGCAGGGGCGCATTAACACTGAACATAGGAAGAAAGAACCCGTTTGTGTCNCCC 2402
DB 4612 AAAAGCAGGGGCGCATTAACACTGAACATAGGAAGAAAGAACCCGTTTGTGTCNCCC 4670
OY 2403 TCGTTGAGGAAGAAATTAATCTGAAAGTCTGGGCAACAGAAAGCAATAGGAGAGCA 2462
DB 4671 TCGTTGAGGAAGAAATTAATCTGAAAGTCTGGGCAACAGAAAGCAATAGGAGAGCA 4729
OY 2463 AAGAAAGCCCGTCTGTGCAAGTTAACTAAAGATTCACCTTCCCTTCCCAAGG 2522
DB 4730 AAGAAAGCCCGTCTGTGCAAGTTAACTAAAGATTCACCTTCCCTTCCCAAGG 4789
OY 2523 CAGTACCCCTCAGACCCCAAGGCCCAACAAGATTCCTAAAGATTTGTAAGACTTAA 2582
DB 4790 CAGTACCCCTCAGACCCCAAGGCCCAACAAGATTTGTAAGACTTAA 4849
OY 2583 GCCCAAGGCTTAATTAACATGATATCTCTGCAATTCCTGAGTGAAGTAAGGA 2642
DB 4850 GCCCAAGGCTTAATTAACATGATATCTCTGCAATTCCTGAGTGAAGTAAGGA 4904
OY 2643 GGGCAAGAAACCAAGTGAAGTGAAGGTTAGTGAAGTCTCAGAGATTAATCAGAG 2702
DB 4905 ATTACAGAAACCAAGTGAAGTGAAGGTTAGTGAAGTCTCAGAGATTAATCAGAG 4962
OY 2703 GCGGTGTCTTTTATACCAAGCTGTAACCTAGCCCTTATACCTGTATTTCCCAATACCA 2762
DB 4963 GCGGTGTCTTTTATACCAAGCTGTAACCTAGCCCTTATACCTGTATTTCCCAATACCA 5022
OY 2763 GAGGAAGCAGAGTGTATACATCTGAGACCTTGAAGATGCTTCTTCTGATCCCTGA 2822
DB 5023 GAGGAAGCAGAGTGTATACATCTGAGACCTTGAAGATGCTTCTTCTGATCCCTGA 5082
OY 2823 CAGCTGACCTCTAATCTTGTGTTGCTTGAAGATCTTCAACCCCACTCACTC 2882
DB 5083 CAGCTGACCTCTAATCTTGTGTTGCTTGAAGATCTTCAACCCCACTCACTC 5142
OY 2883 ACCTGACCTTTTAAACCAAGGTTAGAGGATAGCCCATCTAATTTGGCAGAGATTA 2942
DB 5143 ACCTGACCTTTTAAACCAAGGTTAGAGGATAGCCCATCTAATTTGGCAGAGATTA 5202

OY 2943 GCCCAAGCTTGAAGTATTAATCTGAGCACTTGTCTTCTGTAGTGAAGTAT 3002
DB 5203 GCCCAAGCTTGAAGTATTAATCTGAGCACTTGTCTTCTGTAGTGAAGTAT 5262
OY 3003 TTACTTTTGGCGCTTTCAGAAAACCTTGTGCTTCAAGCCCAACCAAGCTTCTTAA 3062
DB 5263 TTACTTTTGGCGCTTTCAGAAAACCTTGTGCTTCAAGCCCAACCAAGCTTCTTAA 5322
OY 3063 TTCTCGGTAACCTGAGTCAAGGTTTCAAGGATGCTCAAGGCTCAAGTCAAGAG 3122
DB 5323 TTCTCGGTAACCTGAGTCAAGGTTTCAAGGATGCTCAAGGCTCAAGTCAAGAG 5382
OY 3123 GTTAAATCTTGAAGTCAATTAATCAAGGATGCTCAAGGCTCAAGTCAAGAG 3182
DB 5383 GTTAAATCTTGAAGTCAATTAATCAAGGATGCTCAAGGCTCAAGTCAAGAG 5438
OY 3183 CAGCTTATAGGCTTATCTCATCTCAAAACCTTAAAGCACTAAGGATTTCTTGGC 3242
DB 5439 CAGCTTATAGGCTTATCTCATCTCAAAACCTTAAAGCACTAAGGATTTCTTGGC 5498
OY 3243 TTAATAGGTTCTGCGAAMATGATTCGCCAGTGTGCGAATAGCCAGGATTTAA 3302
DB 5499 TTAATAGGTTCTGCGAAMATGATTCGCCAGTGTGCGAATAGCCAGGATTTAA 5557
OY 3303 TACATTAATTAAGAACTCAGAAAGCCATCCCATTTATTAAGTGAAGTGAAG 3362
DB 5558 TACATTAATTAAGAACTCAGAAAGCCATCCCATTTATTAAGTGAAGTGAAG 5617
OY 3363 MRAAGTGTCTTCCAGGCGCCCTTAAAGAGGCTTAAACCAAGTCCCATTTAAAT 3422
DB 5618 MRAAGTGTCTTCCAGGCGCCCTTAAAGAGGCTTAAACCAAGTCCCATTTAAAT 5663
OY 3423 CCAAGGCGCAAGCTTCTTATTAATCAAGAAAGAAAGAAAGAAAGTCTTGAAG 3482
DB 5664 CCAAGGCGCAAGCTTCTTATTAATCAAGAAAGAAAGAAAGTCTTGAAG 5722
OY 3483 CCTTACAGAGTCCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3542
DB 5723 CCTTACAGAGTCCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5782
OY 3543 TGTAGTGAAGAAAGGTTGATCTTATTTTGAAGTGAAGTGAAGTGAAGTGAAG 3602
DB 5783 TGTAGTGAAGAAAGGTTGATCTTATTTTGAAGTGAAGTGAAGTGAAGTGAAG 5842
OY 3603 ATCTGAAGAGTTAAATTAATCAAGGAGAGATCTTATCTGTGAGATCTGATAGT 3662
DB 5843 ATCTGAAGAGTTAAATTAATCAAGGAGAGATCTTATCTGTGAGATCTGATAGT 5902
OY 3663 GAATGAGATTAATCACTGCTTAAAGAGACTTGTGCTGAGACACTGTTTAAATG 3722
DB 5903 GAATGAGATTAATCACTGCTTAAAGAGACTTGTGCTGAGACACTGTTTAAATG 5962
OY 3723 TCAAGCTTATTAATCTGAAGAGGCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 3782
DB 5963 TCAAGCTTATTAATCTGAAGAGGCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 6022
OY 3783 AGTCAATTTCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3842
DB 6023 AGTCAATTTCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6082
OY 3843 TCAACCTATGCACTGAGGAGCTTGTGAAGTGTCTGAGTGTCTGAGTGTCTGAG 3902
DB 6083 TCAACCTATGCACTGAGGAGCTTGTGAAGTGTCTGAGTGTCTGAGTGTCTGAG 6142
OY 3903 ACTTGAATCTGAATGAAGTCTTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3962
DB 6142 ACTTGAATCTGAATGAAGTCTTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6202
OY 3963 TGTGAGTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4022
DB 6202 TGTGAGTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6262
OY 4023 TGCAGAACTAATAGCTTCAATGAGGCACTAAGATTAAGAGAAAGGATTAATA 4082


```
Db      6262 TAGCAGAACTTAATAGGCTTCACTCGGGCACTAGAATTAGAGAAAGGCAAAATA 6321
Qy      4083 TATATACAGACTTBTATATGCTTACCTGCTGTCATGCTCCATGAGCAATATGABAG 4142
Db      6322 TATATACAGACTTBTATATGCTTACCTGCTGTCATGCTCCATGAGCAATATGABAG 6381
Qy      4143 AAAGGAAATTCCTTAATCTTCTGAGAGAACCTATCMACATCAGAAAGCATTAGGARAT 4202
Db      6382 AAAGGAAATTCCTTAATCTTCTGAGAGAACCTATCMACATCAGAAAGCATTAGGARAT 6441
Qy      4203 TATTAATGCGKATGACGAAACCTATAGAGGTGMACTTTACATCTGCTGAGTCAATANA 4262
Db      6442 TATTAATGCGTATGACGAAACCTATAGAGGTGMACTTTACATCTGCGGGTCAATANA 6501
Qy      4263 AAGGAAGRAAGGAAATATASAGRAATGCCAAGCAATATGAAAGMAAAAGAGCAG 4322
Db      6502 AAGGAAGRAAGGAAATATAGAGAGAACTGCCAAGCAATATGAAAGMAAGAGCTG 6561
Qy      4323 CAAGCAGAGACCTTCATTAGAAATGCTTATTAACCTTCCTTATGATAGGGTAATCCCT 4382
Db      6562 CAAGCAGAGACCTTCATTAGAAATGCTTATTAACCAACCCCTAGTATAGGGTAATCCCT 6621
Qy      4383 TCCGGGAAACCAAGCCCGCTACTACGACGAGAAACAGAAATGGGGAACCTCAAGAG-C 4441
Db      6622 TCCGGGAAACCAAGCCCGCTACTACGACGAGAAACAGAAATGGGGAACCTCAAGAGC 6681
Qy      4442 AGTTTCTCCCTCGGAGGCTTAGCACTGAAAGAGGAAATCTTTTGGCTGCACT 4501
Db      6682 AGTTTCTCCCTCGGAGGCTTAGCACTGAAAGAGGAAATCTTTTGGCTGCACT 6741
Qy      4502 ATCCAAATGAAATTAATTAAACCTTCATCAACCTTCACCTTAGGACATGAGACCC 4561
Db      6742 AACCAATGAAATTAATTAAACCTTCATCAACCTTCACCTTAGGACATGAGACCC 6801
Qy      4562 CATCARATGAGCCAAATCATTAATTTACTGACCAAGGCTTTTAAACATATCAAGARATA 4621
Db      6802 CATCARATGAGCCAAATCATTAATTTACTGACCAAGGCTTTTAAACATATTAAGAGATA 6861
Qy      4622 KTCAGAGGCTGTGAATGTGSCARABAAATATCCCTGCTGATCGGCAAGCTCTCA 4681
Db      6862 GTCAGAGGCTGTGAATGTGSCARABAAATATCCCTGCTGATCGGCAAGCTCTCA 6921
Qy      4682 GGAARACAAABAAACAGGCAATTAACCTGABAAABACTGGCAATGATTTTACCAAGC 4741
Db      6922 GGAARACAAABAAACAGGCAATTAACCTGABAAABACTGGCAATGATTTTACCAAGC 6981
Qy      4742 CCAAACTCAGGGAATTCAGTATCTACTAGTCTGGGTATATCTTCAAGGCTTGGGCA 4801
Db      6982 CCAAACTCAGGGAATTCAGTATCTACTAGTCTGGGTATATCTTCAAGGCTTGGGCA 7041
Qy      4802 AGGCTTCCTGCTATGACAGAAAGGCTCAAGAGTATTAAGGCACTAGTCTAGTAA 4861
Db      7042 AGGCTTCCTGCTATGACAGAAAGGCTCAAGAGTATTAAGGCACTAGTCTAGTAA 7101
Qy      4862 TAAATTCAGATTTGGAATTCCTCCGAGGCTTACAGAGTACAAATAGCCCTGCTTCAAG 4921
Db      7102 TAAATTCAGATTTGGAATTCCTCCGAGGCTTACAGAGTACAAATAGCCCTGCTTCAAG 7161
Qy      4922 CCAAGTACCAAGGATTTCCAGAGGCTTACAGAGGCTTACAGATCACTTACATGCGGCT 4981
Db      7162 CCAAGTACCAAGGATTTCCAGAGGCTTACAGAGGCTTACAGATCACTTACATGCGGCT 7221
Qy      4982 GAAAGGCAAGTCTCAGAGGAGGTGAGAAATGAAGAAATCTCAAGAGCACTCA 5041
Db      7222 GAAAGGCAAGTCTCAGAGGAGGTGAGAAATGAAGAAATCTCAAGAGCACTCA 7281
Qy      5042 AAAAGCAAAACCCAGAAACCCACCTCAATGCTGCTGCTGCTTACCTTAA 5101
Db      7282 AAAAGCAAAACCCAGAAACCCACCTCAATGCTGCTGCTGCTTACCTTAA 7341
Qy      5102 GAATCTGCACTTTTCCCAAAAGAGAGCTTACGCTTACGAAATGCTGTATGAGAGC 5161
Db      7342 GAATCTGCACTTTTCCCAAAAGAGAGCTTACGCTTACGAAATGCTGTATGAGAGC 7401
Qy      5162 CCTTATTAACCAATGACCTTGTGCTTGAACCCAGACAGCAACTTATGTTGACATCAC 5221
Db      7402 CCTTATTAACCAATGACCTTGTGCTTGAACCCAGACAGCAACTTATGTTGACATCAC 7460
Qy      5222 CTCCTTACCAAAATATCAACAACTTAAACATTTACAAAGAGACCTATCCCTGAGAA 5281
Db      7461 CTCCTTACCAAAATATCAACAACTTAAACATTTACAAAGAGACCTATCCCTGAGAA 7520
Qy      5282 GGGAAAGAACTATTTCAACCCGCTGACATGATATTAAGTCAAGTCCCTTCTGATATTC 5341
Db      7521 GGGAAAGAACTATTTCAACCCGCTGACATGATATTAAGTCAAGTCCCTTCTGATATTC 7580
Qy      5342 CCCATCCCTAGTACATCTGGGAAAGACCCCTACCAATCACTTTATATATACCCCAATCC 5401
Db      7581 CCCATCCCTAGTACATCTGGGAAAGACCCCTACCAATCACTTTATATACCCCAATCC 7640
Qy      5402 GGTTAAGTGTGCTGAGTGTGATGATCATCACTTGAAGTCAAACTCTGATACT 5461
Db      7641 GGTTAAGTGTGCTGAGTGTGATGATCATCACTGAGTCAAACTCTGATACT 7700
Qy      5462 GCCAAAGAACCTGAAATTCAGAGACACGCTTATCTGTGAACTCTAGAGA 5521
Db      7701 GCCAAAGAACCTGAAATTCAGAGACACGCTTATCTGTGAACTCTAGAGA 7760
Qy      5522 TTTGGGCTGCTCTTCAAAAGCAACAGAGGAAAGTAACTAAATATATAATCCCA 5581
Db      7761 TTTGGGCTGCTCTTCAAAAGCAACAGAGGAAAGTAACTAAATATATAAT- ACCA 7819
Qy      5582 TGAGCTCCCTTATCATATTTTCTKTASTGTSTTTTVAACCTSTTCACTCACTG 5641
Db      7820 TGAGCTCCCTTATCATATTTTCTKTASTGTSTTTTVAACCTSTTCACTCACTG 7879
Qy      5642 CACCCCTCATGCGCTGTATGACAGTACCTCCCTTACMAAGATTCTATGAGAA 5701
Db      7880 CACCCCTCATGCGCTGTATGACAGTACCTCCCTTACMAAGATTCTATGAGAA 7939
Qy      5702 TGAGCGTCCCGGAAATTTGATGCTCCCATGTATAGAGTCTTSTTAAGGAAACCCCA 5761
Db      7940 TGAGCGTCCCGGAAATTTGATGCTCCCATGTATAGAGTCTTSTTAAGGAAACCCCA 7999
Qy      5762 CCTTCACTGCCACACCATATGCCCCGCAACTGCTATCACTGCACTCTTTGCAATG 5821
Db      8000 CCTTCACTGCCACACCATATGCCCCGCAACTGCTATCACTGCACTCTTTGCAATG 8059
Qy      5822 ATGCAATATCTCATTTATGACAGGAAATATGATTAATCTAGTGTCTGAGAGCTTG 5881
Db      8060 ATGCAATATCTCATTTATGACAGGAAATATGATTAATCTAGTGTCTGAGAGCTTG 8119
Qy      5882 GAGTCACTGTCTGTGGAATTTCACTTCAACCAACTGATGCTGATGCGGGGTGAGTTC 5941
Db      8120 GAGTCACTGTCTGTGGAATTTCACTTCAACCAACTGATGCTGATGCGGGGTGAGTTC 8179
Qy      5942 AAGATCAGGCAAGAAACATGTAAGAAAGTATCTCCCACTCAACGSGGTATACATG 6001
Db      8180 AAGATCAGGCAAGAAACATGTAAGAAAGTATCTCCCACTCAACGSGGTATACATG 8239
Qy      6002 GCACCTTACCCCTTACAAAGACATGATTTCTCAAAATCTACATGAAACCTCTGATACC 6061
Db      8240 GCACCTTACCCCTTACAAAGACATGATTTCTCAAAATCTACATGAAACCTCTGATACC 8299
Qy      6062 ATAAGTCTGTGTAAGCTATTTAATTCACACCTCACTGAGGCTTCAAGAGTCTGAGGCC 6121
Db      8300 ATAAGTCTGTGTAAGCTATTTAATTCACACCTCACTGAGGCTTCAAGAGTCTGAGGCC 8359
Qy      6122 AAAACCTTACTTAATGTTGATATGCTCCCTGAACTTACAGGCAATATGTTCAATCC 6181
Db      8360 AAAACCTTACTTAATGTTGATATGCTCCCTGAACTTACAGGCAATATGTTCAATCC 8419
Qy      6182 CTGTACTGAAACAATGGAACAATTTGACGACAGAAATTAACACCACTTCCGTTTATAG 6241
Db      8420 CTGTACTGAAACAATGGAACAATTTGACGACAGAAATTAACACCACTTCCGTTTATAG 8479
```

6242	GACCCCTGGTTCCAAATGGAATAACCCATACCTCAAACTCAAGCTGTGTAATAATT	6301	
6301	GAACCTCTGTTTCCAAATGGAATAACCCATACCTCAAACTCAAGCTGTGTAATAATT	6301	
8480	GAACCTCTGTTTCCAAATGGAATAACCCATACCTCAAACTCAAGCTGTGTAATAATT	8539	
QY	6302	GCAATACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAA	6361
Db	8540	GCAATACTACAGACACAACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAA	8599
QY	6332	TAGTCTGGCTACCCCTCAGAGAAATATTTTGTGTGTGTAACCTCAGCCCTATCGTGTGTA	6421
Db	8600	TAGTCTGGCTACCCCTCAGAGAAATATTTTGTGTGTGTAACCTCAGCCCTATCGTGTGTA	8659
QY	6422	ATGGCTCTTCAGAACTATATGTGCTTCCTCATCTTAACTAGTCCCCCAATGCACTACAA	6481
Db	8660	ATGGCTCTTCAGAACTATATGTGCTTCCTCATCTTAACTAGTCCCCCAATGCACTACAA	8719
QY	6482	CTGAACAAAGATTATACATATATGTCATATCTAAGCCCGGACACAAAAGATACCCATTC	6541
Db	8720	CTGAACAAAGATTATACATATATGTCATATCTAAGCCCGGACACAAAAGATACCCATTC	8779
QY	6542	TTCCCTTTGTTATAGAGACAGAGAGTGCATGAGGACATGATGATGCGATGGACGGATCA	6601
Db	8780	TTCCCTTTGTTATAGAGACAGAGAGTGCATGAGGACATGATGATGCGATGGACGGATCA	8839
QY	6602	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACTGG	6661
Db	8840	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACTGG	8899
QY	6662	TGCGCGATCTCCCTGGTCACTCTGGAAGATCAACTTATCCCTGAGACAGATACCTTC	6721
Db	8900	TGCGCGATCTCCCTGGTCACTCTGGAAGATCAACTTATCCCTGAGACAGATACCTTC	8959
QY	6722	RAAATCGAAGAGCTTTAGACTYGCATACCGCTGAPAGAGGGGAACTGTTTATTTTAC	6781
Db	8960	RAAATCGAAGAGCTTTAGACTYGCATACCGCTGAPAGAGGGGAACTGTTTATTTTAC	9019
QY	6782	GGGAAGAAATGCTGTTATTAATGTTAATCAATCCGGAATGTCATCTGAGAAATTTTAAATA	6841
Db	9020	GGGAAGAAATGCTGTTATTAATGTTAATCAATCCGGAATGTCATCTGAGAAATTTTAAATA	9079
QY	6842	TTCSAGATGGAATPACACGTAAAGAGAGAGACTTCGAAAACACTGGAACCTCGGGGCTCC	6901
Db	9080	TTCSAGATGGAATPACACGTAAAGAGAGAGACTTCGAAAACACTGGAACCTCGGGGCTCC	9139
QY	6902	TCAGCCRAATGATGCGCCTTGATTTCTCCCTCTTAGAGACTCTAGACCTATTAATATTGC	6961
Db	9140	TCAGCCRAATGATGCGCCTTGATTTCTCCCTCTTAGAGACTCTAGACCTATTAATATTGC	9199
QY	6962	TACTCTCTTTTGAGACCCCTGATCTTTBACCTCCTGTTAACTTGTCTTCCAGAAATCG	7021
Db	9200	TACTCTCTTTTGAGACCCCTGATCTTTBACCTCCTGTTAACTTGTCTTCCAGAAATCG	9259
QY	7022	AAGCTGTAAACTACAAATGAGACCCCAAGATCAGTCCAAAGCTAAGATTACCGCAGAC	7081
Db	9260	AAGCTGTAAACTACAAATGAGACCCCAAGATCAGTCCAAAGCTAAGATTACCGCAGAC	9319
QY	7082	CCCTGAGACCGGCTGTATAGCCACAGATCTGAAGTTTAAGACATCAAAAGCACCCCTCTG	7141
Db	9320	CCCTGAGACCGGCTGTATAGCCACAGATCTGAAGTTTAAGACATCAAAAGCACCCCTCTG	9379
QY	7142	AGGAATCTCAGCTGACAACTCTACTACGGCCCAATTACAGACAGAAAGACTTATGAGCG	7201
Db	9380	AGGAATCTTAACTGACAACTCTACTACGGCCCAATTACAGACAGAAAGACTTATGAGCG	9439
QY	7202	GTSGTGCGCAACTCTCCCAACAGACTTAAGTTTCTGTGTAAGTGGGGACCTGAGAG	7261
Db	9440	GTCGTGCGCCAACTCTCCCAACAGACTTAAGTTTCTGTGTAAGTGGGGACCTGAGAG	9499
QY	7262	ACAGGACTAGCTGGAATTTCTTAGGCGATTAAGAACTCTAAGCTCAGTGGGAGAGTGGA	7321
Db	9500	ACAGGACTAGCTGGAATTTCTTAGGCGCGATTAAGAACTCTAAGCTCAGTGGGAGAGTGGA	9559

QY	7322	CCACATCCACCTTTAAACACGGGGCTTGCACTTAGTACACCTGACCAATCAGAGAC	7381
Db	9660	CCACATCCACCTTTAAACACGGGGCTTGCACTTAGTACACCTGACCAATCAGAGAC	9619
QY	7382	TCATTAATATCTTATTTAGGCAAAAGACAGAGGTAAAGAAATAGCAATCATTTATTGCM	7441
Db	9620	TCATTAATATCTTATTTAGGCAAAAGACAGAGGTAAAGAAATAGCAATCATTTATTGCG	9679
QY	7442	TGAGAGCAGCAGCAGAGGGGCAATGATCGGATATATAACCAAGTTCAGAGCGGCAAC	7501
Db	9680	TGAGAGCAGCAGCAGAGGGGCAATGATCGGATATATAACCAAGTTCAGAGCTGGCAAC	9739
QY	7502	GGCAACCCCTTTGGGCTCCCTCCCTTTGTATGGAGCTCTGTTTCATGCTATTTCCT	7561
Db	9740	GGCAACCCCTTTGGGCTCCCTCCCTTTGTATGGAGCTCTGTTTCATGCTATTTCCT	9799
QY	7562	CTATTAAATCTTGACATGCR 7582	
Db	9800	CTATTAAATCTTGACATGCA 9820	
RESULT 15			
LOCUS	AY101587	10229 bp	DNA linear PRI 11-FEB-2004
DEFINITION	Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1 locus, allele B, complete sequence.		
ACCESSION	AY101587		
KEYWORDS	AY101587.1 GI:37544409		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
AUTHORS	1 (bases 1 to 10229) Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Orlol,G., Bonnard,B., Lucotte,G., Duret,L. and Mandrand,B.		
TITLE	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
PUBMED	14757826		
REFERENCE	2 (bases 1 to 10229) Mallet,F., Bouton,O. and Orlol,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142		
PUBMED	CNS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée		
TITLE	d'Italie, Lyon 69364 Cedex 07, France		
JOURNAL	Location/Qualifiers		
FEATURES	1..10229		
source	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/isolate="1"		
	/isolation_source="PBMC of chimpanzee 1"		
	/db_xref="taxon:9598"		
	/cell_type="PBMC"		
	/note="isolated by PCR; ERVW1 locus, allele B		
	endogenous_virus: ERV-W"		
	1..779		
LTR	join(248..810,2889..9820)		
MRNA	/note="putative mRNA transcrip 1; gag-pol pseudogene"		
MRNA	join(248..810,7592..9820)		
	/product="envelope glycoprotein"		
	/note="putative mRNA transcrip 2"		
MRNA	join(248..810,9255..9820)		
	/note="putative mRNA transcrip 3; no coding region"		
CDS	7819..9435		
	/note="syncytin"		
	/codon_start=1		
	/product="envelope glycoprotein"		
	/protein_id="AAM68166.1"		
	/db_xref="GI:37544410"		
	/translation="MALPYHFLFTVLLPSFTLAPPRCMTSSSPFLELMRMRP		
	GNIDPNSRFSKSTPTPTATHTHPRVCVSHATICMANHTYMGKIIINPCGSLG		
	IVCWYTFHTMSDGGGVODARKKHAYEVISQLTRVHSHSRPYKGLDLSGHTLTAT		

HTRLVSLFNTLTGLHEVSAQNPTNCWICLPLNRPVYSIVPEOMNPFSTEINTTSV
LVEPLVSNLEIHTSNTLCVNFSTNTDITDSOCIRWTPPTQIYCLPSGIFVCGTGA
YRCINGSESMCEFLSPVPMITITBEDILNIVYSKRNRPVLPILPVICAGDYLAE
TGIGITSTOPLYKLSOBLNGDMERADSLVLTODLNLAIVLONRALDLITE
RGGTCLFGECCYVUNOSGIYVEKYEIDRIORKEELRNTPMELISQMPWILP
FLRPLAIIILLFPGCTFNULFVNSRIEAVLQMEPKQSKTKLYRRPLDRPASP
RSDVNDIKGPPEIILTAOPLLRNSAGSS"
9494..10229

LTR
Query Match 83.7%; Score 6344.8; DB 9; Length 10229;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 6624; Conservative 194; Mismatches 111; Indels 52; Gaps 28;

QY 606 GATGGGAAAGCTTCCCGCAAGACAAAAGCCCTTAAGACGATTCGABAAATGGGA 665
DB 2888 GATGGGAAAGCTTCCCGCAAGACAAAAGCCCTTAAGACGATTCGABAAATGGGA 2947
QY 666 MCAATTGACCTCTCAGACACTAAGAAAGAAAGCACTTATATCTTCGACGTCGCGCTG 725
DB 2948 CCAATTGACCTCTCAGACACTAAGAAAGAAAGCACTTATATCTTCGACGTCGCGCTG 3007
QY 726 GCACCTCTAGAGGAAGTAAATTAATTAACCACTTACAGCTAGACCTTTTGTGA 785
DB 3008 GCACCTCTAGAGGAAGTAAATTAATTAACCACTTACAGCTAGACCTTTTGTGA 3067
QY 786 AAGGCAATAGAGTGAAGTGCCTAATAGTACAACTTCTTTCATTAAGAGCACTCA 845
DB 3068 AAGGCAATAGAGTGAAGTGCCTAATAGTACAACTTCTTTCATTAAGAGCACTCA 3127
QY 846 AATTATGTAAGAAAGTGTATTTATGCTTACAGAAAGCTTCAGAGTCACTCCCTATC 905
DB 3128 AATTATGTAAGAAAGTGTATTTATGCTTACAGAAAGCTTCAGAGTCACTCCCTATC 3187
QY 906 CGAGATCCCC--GACTCCTCCCGCAMTAAATTAAGACCCCTTCAACCAATGTGTC 963
DB 3188 CGAGATCCCCGACTGACTCCTCCCGCAMTAAATTAAGACCCCTTCAACCAATGTGTC 3247
QY 964 AAAAGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCTAATTCGCCAATTAT 1023
DB 3248 AAAAGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCTAATTCGCCAATTAT 3307
QY 1024 GACCCCTCCCAAGCAGTGGAGAGAGAAATTCGCCCCAGCAGATGTCATGCTTTT 1083
DB 3308 GACCCCT--CAAAGCAGTGGAGAGAGAAATTCGCCCCAGCAGATGTCATGCTTTT 3366
QY 1084 YYTCTCCAGACTTAAGCAATTAAGAAATTAAGCACTTAAGTAAATTCAGATAATCTGAT 1143
DB 3367 TCTCTCAGACTTAAGCAATTAAGAAATTAAGCACTTAAGTAAATTCAGATAATCTGAT 3426
QY 1144 GGCATATTTGATGTTTACAGAGGTTAGGACAAATCTTTCATCTGACATGAGAGATATA 1203
DB 3427 GGCATATTTGATGTTTACAGAGGTTAGGACAAATCTTTCATCTGACATGAGAGATATA 3486
QY 1204 TATGTCATGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACCATTAATCTGAC 1263
DB 3487 -ATGTCATGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACCATTAATCTGAC 3545
QY 1264 CTGAGRTTGGCGATCTCTGGTATCTCAGTACAGTCAATGATNNGATGACACAGAA 1323
DB 3546 CCAAGAGTTTGGCGATCTCTGGTATCTCAGTACAGTCAATG--ATAGATGACACAG-A 3602
QY 1324 GGAAGAAATATGATCCCGCAGCGCAGCAGCACTTCCAGTCTASACCCCATTTGGGG 1383
DB 3603 GGAAGAAATATGATCCCGCAGCGCAGCAGCACTTCCAGTCTASACCCCATTTGGGG- 3661
QY 1384 ACAAGAAATATGATCACTGAGAGATGTGTCTGACACTTTGCTAACTTGTGTCTAS 1443
DB 3662 --ACAAGAAATATGATCACT--GAGATGTGTCTGACACTTTGCTAACTTGTGTCTAG 3728
QY 1444 AAGGACTTAAGGAAATCTASGAAAGAAATCTTAAGATTACTCAATGATCTCAACCTAACA 1503
DB 3719 AAGGACTTAAGGAAATCTAGAGAG--AAGCTTATGAATTACTCAATGATCTCAACCTAACA 3777

QY 1504 CAGGGGAAGGAAAGAAATCTTAATCTGCTTTTCTGAGAGACTAAGGAGCATTTAGAA 1563
DB 3778 CA-GGGAAAGGAAAGAAATCTTAATCTGCTTTTCTGAGAGACTAAGGAGCATTTAGAA 3836
QY 1564 GCGTGGCTCTGTGTCACCTGACCTCTGAAAGCCCACTAATCTTAAAGGCTAAGTTAT 1623
DB 3837 GCGTGGCTCTGTGTCACCTGACCTCTGAAAGCCCACTAATCTTAAAGGCTAAGTTAT 3896
QY 1624 CACTAGTCAGCTGACAGACTTAAAG-AAAAAATCTTAAAGTCTGCGTAGGCCGAGCA 1682
DB 3897 CACTAGTCAGCTGACAGACTTAAAGAAAAAATCTTAAAGTCTGCGTAGGCCGAGCA 3956
QY 1683 AAACTTAGAAACCTTATTAATCTTGGCACTTCGCTTTTATTAATAGATCAAGAGAA 1742
DB 3957 AAACTTAGAAACCTTATTAATCTTGGCACTTCGCTTTTATTAATAGATCAAGAGAA 4016
QY 1743 GCAGCGGAAACAGACAAACGGAATTAATAAAGGCAACGCTTAAATGATGACCTCA 1802
DB 4017 GCAGCGGAAACAGACAAACGGAATTAATAAAGGCAACGCTTAAATGATGACCTCA 4076
QY 1803 GCAAGTGAATTTTGAAGCTCTGGAAGAAAGGAAAGCTGGCAATTGAATGCTTAATA 1862
DB 4077 GCAAGTGAATTTTGAAGCTCTGGAAGAAAGGAAAGCTGGCAATTGAATGCTTAATA 4136
QY 1863 GGGCTTGTCTCAGTGCAGTCTCAAGAGCACTTTAAAAAAGATTGTCCAATGAGATA 1922
DB 4137 GGGCTTGTCTCAGTGCAGTCTCAAGAGCACTTTAAAAAAGATTGTCCAATGAGATA 4196
QY 1923 AGCCGCCCTTGTGTCAGTCCCTTAATTAAGGAAATCACTGGAAGGCCACTGGCCCA 1982
DB 4197 AGCCGCCCTTGTGTCAGTCCCTTAATTAAGGAAATCACTGGAAGGCCACTGGCCCA 4256
QY 1983 GGGGCAAAAGTCTTTTGAATGACAGAGCACTAACCAATGATCAAGACAGACAGACTGAG 2042
DB 4257 GGGGCAAAAGTCTTCTGATGACAGAGCACTAACCAATGATCAAGACAGACAGACTGAG 4316
QY 2043 GGTGCTGGGGCAAGCGCATCCCATGTCATCACTTCAAGAGGCCCTGGGTATGCTTGA 2102
DB 4317 GGTGCTGGGGCAAGCGCATCCCATGTCATCACTTCAAGAGGCCCTGGGTATGCTTGA 4376
QY 2103 CCATTGAAGGCGCAGGAAAGTGTCTCTGAGACACTGAGTGGCGTCTTCTTAAGCTTAATCT 2162
DB 4377 CCATTGAAGGCGCAGG-AGTGTGTCTCTGAGACACTGAGTGGCGTCTTCTTAAGCTTAATCT 4435
QY 2163 TCTGTCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCNTTAAGACGG 2222
DB 4436 TCTGTCCCGGACAACTGTCTCTCAGATCTGTCACTA-TCTGAGGGGTC--TAAAGACGG 4493
QY 2223 CAGTCACTAAGTATCTTTTCCCAAGCACTAAGTATGAATCGGGAGCTTATTTCTTTTC 2282
DB 4494 CAGTCACTAAGTATCTTTTCCCAAGCACTAAGTATGAATCGGGAGCTTATTTCTTTTC 4551
QY 2283 ACATGCTTTTCTAATTAATGCTTGAAGAGCCCACTACCTTGTGAAGAGAGACATTCCTGAC 2342
DB 4552 ACATGCTTTTCTAATTAATGCTTGAAGAGCCCACTACCTTGTGAAGAGAGACATTCCTGAC 4611
QY 2343 AAAAGCAGGGGCACTTAATACCTGAACTAAGAGAAAGAACACCGGTTGTGTNCCC 2402
DB 4612 AAAAGCAGGGGCACTTAATACCTGAACTAAGAGAAAGAACACCGGTTGTGTNCCC 4670
QY 2403 TGCCTTAGAGAAAGAAATTAATCTTGAATCTGGGCAAGAAAGCAATATGAGAGAGCA 2462
DB 4671 TGCCTTAGAGAAAGAAATTAATCTTGAATCTGGGCAAGAAAGCAATATGAGAGAG-CA 4729
QY 2463 AAGAAATCCCGTCTGTCTCAAGTTAACTAAAGATTCACCTTCCCTTCAACAAAG 2522
DB 4730 AAGAAATCCCATCTGTCTCAAGTTAACTAAAGATTCACCTTCCCTTCAACAAAG 4789
QY 2523 CAGTACCCCTCAGACCCCAAGGCCCAACAAAGATTCACAAAGATTTGTAAGACTTAATA 2582
DB 4790 CAGTACCCCTCAGACCCCAAGGCCCAACAAAGATTCACAAAGATTTGTAAGACTTAATA 4849

2583 GCCCAGGCTTAAATACCATGATTAACCTCCCTGACGTAATCCGATGATGAGGA 2642
Db GCCCAGGCTTAAATACCATGATTAACCTCCCTGACGTAATCCGATGATGAGGA 4804
2643 GGCACAGAAAACCGATGAGCAGTGAAGGTTAGTGCAGATCTCAGGATTAATGAG 2702
Db ATTACAGAAAACCGATGAGCAGTGAAGGTTAGTGCAGATCTCAGGATTAATGAG 4905
2703 GCCGTTGCTTTTAAACCCAGCTGTAACCTGACCTTAACTGTGATTTCCAAATGCA 2762
Db GCGTTGTTCTCTAATAGCCAGCTGATCAGCCCTTAATCTGCTTTCCAAATGCA 4963
2763 GAGGAAGCAGAGGTTTAACTGACCTGACCTTAAAGATGCTTTCTGACATCCCTGTA 2822
Db GAGGAAGCAGAGGTTTAACTGACCTGACCTTAAAGATGCTTTCTGACATCCCTGTA 5023
2823 CATCTGACCTCTCAATTTCTGTTTGCTTTGAGATGATCTTCAACCCAGATCTCAATC 2882
Db CATCTGACCTCTCAATTTCTGTTTGCTTTGAGATGATCTTCAACCCAGATCTCAATC 5083
2883 ACCTGACCTTTTAAACCCAGGTTTCAAGGATAGGCCCATCTATTGCGCAGCATTA 2942
Db ACTGACATATTTTAAACCCAGGTTTCAAGGATAGGCCCATCTATTGCGCAGCATTA 5143
2943 GCCCAAGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3002
Db GCCCAAGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5203
3003 TTACTTTTGGCGCTTGAAGAACTTGGCATGACGACCCAGGCTCTTAAAT 3062
Db TTACTTTTGGCGCTTGAAGAACTTGGCATGACGACCCAGGCTCTTAAAT 5263
3063 TTCCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3122
Db TTCCTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5323
3123 GTTAAATCTTAAGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3182
Db GTTAAATCTTAAGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 5383
3183 GACCTATATCTGCTTATCTCATCTCAATCAATCAATCAATCAATCAATCAATCAAT 3242
Db GACCTATATCTGCTTATCTCATCTCAATCAATCAATCAATCAATCAATCAATCAAT 5439
3243 RTAAATAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3302
Db RTAAATAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5499
3303 TACATTAATTAAGAACTGAAAGGCAATCCATTTAATTAATTAATTAATTAATTAAT 3362
Db TACATTAATTAAGAACTGAAAGGCAATCCATTTAATTAATTAATTAATTAATTAAT 5558
3363 MRATAGGCTTCTGAGGCTTAAAGAGGCTTAAAGAGGCTTAAAGAGGCTTAAAGAG 3422
Db MRATAGGCTTCTGAGGCTTAAAGAGGCTTAAAGAGGCTTAAAGAGGCTTAAAGAG 5618
3423 CCAACGAGGCAAGCTTTTSTTATATATCAAGAAAAAACAAGAAATGCTGAGAT 3482
Db CCAACGAGGCAAGCTTTTSTTATATATCAAGAAAAAACAAGAAATGCTGAGAT 5664
3483 CCTTACAGATCCRAAGAGAGCTTGAACCTRTGCGCACTGATGATGATGATGATGAT 3542
Db CCTTACAGATCCRAAGAGAGCTTGAACCTRTGCGCACTGATGATGATGATGATGAT 5723
3543 TGTAGTGGCAAGGCTTGTCTTGTATGATGATGATGATGATGATGATGATGATGAT 3602
Db TGTAGTGGCAAGGCTTGTCTTGTATGATGATGATGATGATGATGATGATGATGAT 5783
3603 ATCTGAAGCAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3662
Db ATCTGAAGCAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5843
3663 GAATGCAATCTCACTGCTAAAGAGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3722

5903 GAATGCAATCTCACTGCTAAAGAGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5962
3723 TCAAGCTTATTAATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3782
Db TCAAGCTTATTAATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5963
3783 AGTCATTTCTTCAAGCAATGAAGAAAGATTAATTAATTAATTAATTAATTAATTAAT 3842
Db AGTCATTTCTTCAAGCAATGAAGAAAGATTAATTAATTAATTAATTAATTAATTAAT 6023
3843 TCATACTATGCACTGAGAGGCACTTATGAAGTTCCYTGAATGATTCGACCTTCA 3902
Db TCATACTATGCACTGAGAGGCACTTATGAAGTTCCYTGAATGATTCGACCTTCA 6083
3903 ACTTGAATCTGATGAGAGTTCTTTGTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3962
Db ACTTGAATCTGATGAGAGTTCTTTGTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 6142
3963 TGTGATGATTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4022
Db TGTGATGATTAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6202
4023 TGCAGAACTAATGAGCTTCAATGAGGCACTAAGATTAAGAGAGAGAGAGAGAGAGAG 4082
Db TGCAGAACTAATGAGCTTCAATGAGGCACTAAGATTAAGAGAGAGAGAGAGAGAGAG 6262
4083 TATATCAGACTTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4142
Db TATATCAGACTTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 6322
4143 AAAGGAAATCTTAACTTCTGAGAGAACTATCAATCAATCAATCAATCAATCAATCAAT 4202
Db AAAGGAAATCTTAACTTCTGAGAGAACTATCAATCAATCAATCAATCAATCAATCAAT 6382
4203 TATATGAGCTGATGAGAACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4262
Db TATATGAGCTGATGAGAACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6442
4263 AAGGAAGAAAGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4322
Db AAGGAAGAAAGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6502
4323 CAAGGAG 4382
Db CAAGGAG 6562
4383 TCCGGAAGAAACCAAGCCAGTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4441
Db TCCGGAAGAAACCAAGCCAGTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6622
4442 AGTTTCTCTCTGAG 4501
Db AGTTTCTCTCTGAG 6682
4502 ATCCATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4561
Db ATCCATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6742
4562 CATCATATGAGCAATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4621
Db CATCATATGAGCAATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6802
4622 KTCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4681
Db KTCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6862
4682 GGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4741
Db GGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6922
4742 CCAAACTCAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4801

Dh 6982 CCAATCTCAGGGATTTCAGTATCTACTAGTCTGGGTAGATATCTTTCACGGGTTGGCAG 7041
Qy 4802 AGGCTTCCCTGTAGAGCAGAAAAAGGCCNAAGGTAAATPAAGSCATAGTTGATGAA 4881
Dh 7042 AGGCTTCCCTGTAGAGCAGAAAAAGGCCNAAGGTAAATPAAGSCATAGTTGATGAA 7101
Qy 4862 TAAATCCAGATTTCGACCTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG 4921
Dh 7102 TAAATCCAGATTTCGACCTTCCCGAGGCTTACAGATGACAAAGCCCTGCTTCAAG 7161
Qy 4922 CCAAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCT 4981
Dh 7162 CCAAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCT 7221
Qy 4982 GAAAGCCACAGTCTCAGGGAAGGTGAGAAAAATGAATGAATACTGAAGGACATCTAA 5041
Dh 7222 GAAAGCCACAGTCTCAGGGAAGGTGAGAAAAATGAATGAATACTGAAGGACATCTAA 7281
Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCAATGCGCTGTCTGTGCTTATAGCCTTAAAA 5101
Dh 7282 AAAAGCAAAACCCAGGAAACCCACCTCAATGCGCTGTCTGTGCTTATAGCCTTAAAA 7341
Qy 5102 GAATCGCAACTTCCCGCAAAAAGAGSACTTACGAAATGCTTACGAAATGCTTATGAAAGC 5161
Dh 7342 GAATCGCAACTTCCCGCAAAAAGAGSACTTACGAAATGCTTATGAAAGC 7401
Qy 5162 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCAACTTATGTTGACAGATCAC 5221
Dh 7402 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCAACTTATGTTGACAGATCAC 7460
Qy 5222 CTCTTACGCCAAATATCAACAAGTTCTTAAACATTAACAAGAACCTATCCCTGAGAAG 5281
Dh 7461 CTCTTACGCCAAATATCAACAAGTTCTTAAACATTAACAAGAACCTATCCCTGAGAAG 7520
Qy 5282 GGGAAAGAACTATTCCACCCMWGTGACATGTAATTAATGTAAGTCCCTTCTCTTAATTC 5341
Dh 7521 GGGAAAGAACTATTCCACCCMWGTGACATGTAATTAATGTAAGTCCCTTCTCTTAATTC 7580
Qy 5342 CCCATCCCTAGATACATCTCTGGGAAGACCTTACCCAGTCACTTTATTAATACCCCACTGC 5401
Dh 7581 CCCATCCCTAGATACATCTCTGGGAAGACCTTACCCAGTCACTTTATTAATACCCCACTGC 7640
Qy 5402 GGTAAAGTGGCTGAGTGAAGTCTTGAATACATCACTTGAAGTGAATCCCTGATACT 5461
Dh 7641 GGTAAAGTGGCTGAGTGAAGTCTTGAATACATCACTTGAAGTGAATCCCTGATACT 7700
Qy 5462 GCCAAAGAACTGAAAAATCCAGAGACAAAGCTTACTTCTGTTGAACCTCTAGAGAA 5521
Dh 7701 GCCAAAGAACTGAAAAATCCAGAGACAAAGCTTACTTCTGTTGAACCTCTAGAGAA 7760
Qy 5522 TTTGGCGCTGCTTCAACCAACACAGAGAAAGTAACCTAAATCATTAATCCGCCA 5581
Dh 7761 TTTGGCGCTGCTTCAACCAACACAGAGAAAGTAACCTAAATCATTAAT-ACCCA 7819
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCSTTTTCACTCTCATCG 5641
Dh 7820 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCSTTTTCACTCTCATCG 7879
Qy 5642 CACCCCTTCATATGCGGCTATAGACAGTATCCCTTACCMAGAGTTTCTATGAGAA 5701
Dh 7880 CACCCCTTCATATGCGGCTATAGACAGTATCCCTTACCMAGAGTTTCTATGAGAA 7939
Qy 5702 TGACGCTCCGGAATAATGATGCCCCATCGTATAGAGTCTTTSTAAGGAAACCCCA 5761
Dh 7940 TGACGCTCCGGAATAATGATGCCCCATCGTATAGAGTCTTTSTAAGGAAACCCCA 7999
Qy 5762 CTTTACATGCCCAACCCATATAGCCCGCACTGTATCATCTGCACTCTTTCATATGC 5821
Dh 8000 CTTTACATGCCCAACCCATATAGCCCGCACTGTATCATCTGCACTCTTTCATATGC 8059
Qy 5822 ATGCAAAATCTATTTATGGAAGAAAAATGATTAATCTTATGTTGCTGAGAGACTTG 5881
Dh 8060 ATGCAAAATCTATTTATGGAAGAAAAATGATTAATCTTATGTTGCTGAGAGACTTG 8119

Qy 5882 GAGTCACTGCTCTTGTGACTTACTTACCCAACTGATATGCTGATGGGGTGAAGTTC 5941
Dh 8120 GAGTCACTGCTCTTGTGACTTACTTACCCAACTGATATGCTGATGGGGTGAAGTTC 8179
Qy 5942 AAGATCAGGCAAGAGAAAAATGTAAAGAAATTAATCTCCAACTCAGCCGGGTACATG 6001
Dh 8180 AAGATCAGGCAAGAGAAAAATGTAAAGAAATTAATCTCCAACTCAGCCGGGTACATG 8239
Qy 6002 GCACTCTTACCCCTTACAAAGACATGATCTTCAAAATCATAGAAACCTCCGTATACC 6061
Dh 8240 GCGCTCTTACCCCTTACAAAGACATGATCTTCAAAATCATAGAAACCTCCGTATACC 8299
Qy 6062 ATATCTGCTGCTTAAAGCCCTTATTAACACACCTTCACTGGGCTCATGAGTCTCGGCC 6121
Dh 8300 ATATCTGCTGCTTAAAGCCCTTATTAACACACCTTCACTGGGCTCATGAGTCTCGGCC 8359
Qy 6122 AAAACCTTACTAAGTTGATATGCTCCCTGAACTTCARAGCATATGTTCAATCC 6181
Dh 8360 AAAACCTTACTAAGTTGATATGCTCCCTGAACTTCARAGCATATGTTCAATCC 8419
Qy 6182 CTGTACTGAAACATGGAACAACTTCAGACAGAAATTAACACCACTTCCGTTTATAG 6241
Dh 8420 CTGTACTGAAACATGGAACAACTTCAGACAGAAATTAACACCACTTCCGTTTATAG 8479
Qy 6242 GACCTCTTGTTCGAATSTGGAATTAACCATACCTCAACCTGCTGTAATTA 6301
Dh 8480 GACCTCTTGTTCGAATSTGGAATTAACCATACCTCAACCTGCTGTAATTA 8539
Qy 6302 GCAATACTACATACAAACCACTCCCAATGATGAGTGGGTAACTCCCTCCACACAA 6361
Dh 8540 GCAATACTACATACAAACCACTCCCAATGATGAGTGGGTAACTCCCTCCACACAA 8599
Qy 6362 TAGTCTGCTTACCTTCAAGAAATTTTGTGTGTGTGTAACCTGAGCTTATGTTTGA 6421
Dh 8600 TAGTCTGCTTACCTTCAAGAAATTTTGTGTGTGTGTAACCTGAGCTTATGTTTGA 8659
Qy 6422 ATGCTCTTCAAGAACTTATGCTGCTTCACTTATGCTGCTTCACTTATGCTGCTTCA 6481
Dh 8660 ATGCTCTTCAAGAACTTATGCTGCTTCACTTATGCTGCTTCACTTATGCTGCTTCA 8719
Qy 6482 CTGAACAAGTTTATACAGTTATGTCATATCTTAAACCCCGCAACAAAGAGTACCATTC 6541
Dh 8720 CTGAACAAGTTTATACAGTTATGTCATATCTTAAACCCCGCAACAAAGAGTACCATTC 8779
Qy 6542 TTTCTTTTGTATAGAGACAGAGTCTAGGTGACATAGTATCTGACATGCGGTATCA 6601
Dh 8780 TTTCTTTTGTATAGAGACAGAGTCTAGGTGACATAGTATCTGACATGCGGTATCA 8839
Qy 6602 CAACCTTACTACAGTTCTACTACAAACATATCTCAAGAACTTAAATGGGACATGGAACGG 6661
Dh 8840 CAACCTTACTACAGTTCTACTACAAACATATCTCAAGAACTTAAATGGGACATGGAACGG 8899
Qy 6662 TCGCGACTCCCTGCTCACTTGGCAAGATCAATTAATCCCTTACAGACAGTATGCTTC 6721
Dh 8900 TCGCGACTCCCTGCTCACTTGGCAAGATCAATTAATCCCTTACAGACAGTATGCTTC 8959
Qy 6722 TAAATCGAAGCTTATAGACTTGTCTTACACCGCTGABAGAGGGGAACTGTTATTTTATG 6781
Dh 8960 TAAATCGAAGCTTATAGACTTGTCTTACACCGCTGABAGAGGGGAACTGTTATTTTATG 9019
Qy 6782 GGGAAAGATGCTGTTATTAATGTTATCAATCCGGAATGCTACATGAGAAAGTTTAAAGAA 6841
Dh 9020 GGGAAAGATGCTGTTATTAATGTTATCAATCCGGAATGCTACATGAGAAAGTTTAAAGAA 9079
Qy 6842 TTSAGATCGAATAACAAGTAKAGACAGAGGCTTTCGAAACATGAGACCTGAGGCTCC 6901
Dh 9080 TTSAGATCGAATAACAAGTAKAGACAGAGGCTTTCGAAACATGAGACCTGAGGCTCC 9139
Qy 6902 TCAAGCAATGGAATGCTGAGTCTCCCTTCTTATAGACCTTATAGACCTTATATTTGC 6961
Dh 9140 TCAAGCAATGGAATGCTGAGTCTCCCTTCTTATAGACCTTATAGACCTTATATTTGC 9199

Db	4910	CCAGATCCCCGACTCTTCTCCCACTAATAAGACCCCCCTTCAACCATAATGGTCCAA	4969
Qy	966	AAGAGATAGACAAAAGGGTAAACAGTAAACCAAGAGTGGCAATATTCCCAATTATGA	1025
Db	4970	AAGGAGATAGACAAAAGGGTAAACAGTAAACCAAGAGTGGCAATATTCCCAATTATGA	5029
Qy	1026	CCCCCTCCCAAGCAGTGGGAGAGAAATTGGCCACAGAGTGCATGTCTTTTYY	1085
Db	5030	CCCCCTCCCAAGCAGTGGGAGAGAAATTGGCCACAGAGTGCATGTCTTTTTC	5088
Qy	1086	TCTCCCAAGCTTAAAGCAATTAATAACAGACTTAGTAAATTCTCAGTAAACTTGAATGG	1145
Db	5089	TCTCCCAAGCTTAAAGCAATTAATAACAGACTTAGTAAATTCTCAGTAAACTTGAATGG	5148
Qy	1146	CTAATTGTGTGTTTAAACAAGGGTTAGCAATCTTTGATCTGACATGAGAGATATATA	1205
Db	5149	CTAATTGTGTGTTTAAACAAGGGTTAGCAATCTTTGATCTGACATGAGAGATATTA-A	5207
Qy	1206	TGTCACCTGCTAATCAGACACTTAAACCCCAATTGAGAGAGTGCACATACTGACGCT	1265
Db	5208	TGTCACCTGCTAATCAGACACTTAAACCCCAATTGAGAGAGTGCACATACTGACGCT	5267
Qy	1266	GAGRGTTTGGCGATCTTGTTATCTCACTCAGTCACTGATGATANGATGACACAGAAAG	1325
Db	5268	GAGAGTTTGGCGATCTTGTTATCTCACTCAGTCACTGATGATANGATGACACAGAAAG	5324
Qy	1326	AAAGAAATGATTTCCCAACAGGCCACAGCAGATGCCAGTCTASACCTCATTTGGGGAC	1385
Db	5325	AAAGAAATGATTTCCCAACAGGCCACAGCAGATGCCAGTCTCAACCTCATTTGGG--	5381
Qy	1386	ACAGAAATCAGTAAACATGAGGAGATTGTGTGCTCAGACATTTGCTAACTTGTGTCTASAA	1445
Db	5382	ACACAGAAATCAGAAACAT-AGAAATGTGTGTGCTCAGACATTTGCTAACTTGTGTCTASAA	5440
Qy	1446	GGACTAAGGAAAACTASGAAAGAAATCTAYGAATTACTCATATGATGTCCACATTAACA	1505
Db	5441	GGACTAAGGAAAACTASGAAAG--AAGTATAGAAATTAACATAGATGTCCACATTAACA	5499
Qy	1506	GGGGAAGGGAAGAAAAATCCTACTGCGCTTTGAGAGAGACTAAGGAGGCAATTGAGGAAGC	1565
Db	5500	-GGGAAGGGAAGAAAAATCCTACTGCGCTTTGAGAGAGACTAAGGAGGCAATTGAGGAAGC	5558
Qy	1566	GTGCTCTGTCTCACCTGACTCTTCTGAAAGCCCACTAATCTTAAAGCGTAAGTTATCA	1625
Db	5559	GTGCTCTGTCTCACCTGACTCTTCTGAAAGCCCACTAATCTTAAAGCGTAAGTTATCA	5618
Qy	1626	CTCAGTCAGCTGACAGCATTAAG-AAAAAATTCAAAAGTCTGCGCTAGGCGCGAGCAAA	1684
Db	5619	CTCAGTCAGCTGACAGCATTAAGAAAAAACTTCAAAAGTCTGCGCTAGGCGCGAGCAAA	5678
Qy	1685	ACTTAGAAAACTTATTTGAACCTTGGCAACTTGGTTTTTATATATAGATCAGAGGGAGC	1744
Db	5679	ACTTAGAAAACTTATTTGAACCTTGGCAACTTGGTTTTTATATATAGATCAGAGGGAGC	5738
Qy	1745	AGGCGGAACAGGACAAACGGGATTAATAAAAAAGGCCACCGCTTAGTCATGACCTCAGG	1804
Db	5739	AGGCGGAACAGGACAAACGGGATTAATAAAAAAGGCCACCGCTTAGTCATGACCTCAGG	5798
Qy	1805	CAAGTGAATTTTGAAGGCTCTGGAAAAAGGAAAAAGCTGGGCAATTTGAATGCTAATAGG	1864
Db	5799	CAAGTGAATTTTGAAGGCTCTGGAAAAAGGAAAAAGCTGGGCAATTTGAATGCTAATAGG	5858
Qy	1865	GCTTGCTTCCAGTGGCGGTCTACAGGACCTTTAAAAAAAGTTTTCAGATGAGAGTAAAG	1924
Db	5859	GCTTGCTTCCAGTGGCGGTCTACAGGACCTTTAAAAAAAGTTTTCAGATGAGAGTAAAG	5918
Qy	1925	CGGCCCCCTTCCTCATGCCCCCTTAATTTCAAGGAATCACTGGAAAGGCCACCTGCCCCAGG	1984
Db	5919	CGGCCCCCTTCCTCATGCCCCCTTAATTTCAAGGAATCACTGGAAAGGCCACCTGCCCCAGG	5978
Qy	1985	GGACAAAGGCTTTTAGTCAGAAAGCACTTAACAGATGATCCAGCAGGACTTGAAGG	2044

D	579	GGACAAAGGTCCTCTGATGTCAGAAAGCCATTAACNAATATATCCAGACGAGACTGAGGG	6038
Q	2045	TGCTCTGGGGGCAAGCGCCATCCATGCCATCACTCTCAACAGAGCCCTGGGTATGCTTGACC	2104
D	6039	TGCTCTGGGGGCAAGCGCCATCCATGCCATCACTCTCAACAGAGCCCTGGGTATGCTTGACC	6098
Q	2105	ATTGAGGGGCAAGGAAGTTGTCTCCAGGACACTGGAGCGAGTCTTCTTAAGCTTAACCTTC	2164
D	6099	ATTGAGGGGCAAGG-AAGTTGTCTCTGAGACACTGGAGCGAGTCTTCTTAAGCTTAACCTTC	6157
Q	2165	TGTCGCCGAGCAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGCGGGCA	2224
D	6158	TGTCGCCGAGCAACTGTCTCTCAGATCTGTCACTA-ACGTGAGGGGGTCC-TAAACGGGGCA	6215
Q	2225	GTCATAGAAATCTTTATCCGAGCACTAAGTTAATGAACGGGGAGCTTTATCTTTTCAC	2284
D	6216	GTCATAGAAATC-TTCTCCAGAGCACTAAGTTATG-ACTGGGGAGCTTTATCTTTTCAC	6273
Q	2285	ATGCTTTTCTAATTAATGCTTTGAAGGCCCACTACCTTGTTTGGGAGAGACATTCTAGCAA	2344
D	6274	ATGCTTTTCTAATTAATGCTTTGAAGGCCCACTACCTTGTTTAGGAGAGACATTCTAGCAA	6333
Q	2345	AAGCAGGGGCACTTATTAACCTTAACAATAGAGAAAGAAACCCCGTTGTGTGTTGCCCTG	2404
D	6334	AAGCAGGGGCACTTATTAACACTTAACAATAGAGAAAGAAACCCCGTTGTGTG-CCCTGG	6392
Q	2405	CTTGAGGAGAAATTAATCCTGAAGTCTGGGCAACAGAAAGAACATAATGAGCAGCCAAA	2464
D	6393	CTTGAGGAGAAATTAATCCTGAAGTCTGGGCAACAGAAAGAACATAATGAGCAG-CAAA	6451
Q	2465	GAATGCCCGTCTGTTCAGATTAACTAAAGATTCTACTTTCCCTTCCCAAGAGGCA	2524
D	6452	GAATGCCCGTCTGTTCAGATTAACTAAAGATTCTCACTCTTTCCCTTCCCAAGAGGCA	6511
Q	2525	GTAACCCCTCAAGACCCAGAGCCCAACAAAGATTCTCAAAAATTTGTTAAGACTTAAAGC	2584
D	6512	GTAACCCCTCAAGACCCAGAGCCCAACAAAGACTTCAAAAATTTGTTAAGACTTAAAGC	6571
Q	2585	CCAAAGGCTGTGTAATAACCATGACATAACTCCCTSCAGTAATTCGTAGTGAAGAGG	2644
D	6572	CCAAAGGCTGTGTAATAACCATGACATAACCTCTCAGTAACCTC-----AATTTTAGAGG	6624
Q	2645	CACAGAAACCCAGTGGACATGTGAGGGTTGTGTCACAGATCTCAGATTAACAATGAGGC	2704
D	6625	TACAGAAACCCCAACAGACAGTGGAGGTATGTGACAGATCTCAGATTAACAAT-GAGGC	6682
Q	2705	CGTTGCTCTTATTAACCCAGCTGACCTAGCCCTTAATCTAGTGTGTTCCCAATATCCAGA	2764
D	6683	TGTTGTCTCTTAATAGCAGACTGACCTAGCCCTTAATCTAGTGTGTTCCCAATATCCAGA	6742
Q	2765	GGAAAGCAGAGTGGTTTACASTCTGGAACCTTMAAGATGCTTCTTCTGCAATCCCTGACA	2824
D	6743	GGAAAGCAGAGTGGTTTACAGTCCGGAACCTTCAAGAGGCTTCTTCTGCAATCCCTGACA	6802
Q	2825	TCCGACATCTCAATCTTGTGTTGGCTTTGAAGATATCTTCAACCCCAACATCTCAACTCAC	2884
D	6803	TCCGACATCTCAATCTTGTGTTGGCTTTGAAGATATCTTCAACCCCAACATCTCAACTCAC	6862
Q	2885	CTGGAATRTTTTACCCCAAGGGTTCAAGGATATAGTCCCACTATTTTGGCCAGGCATTAGC	2944
D	6863	CTGGAATRTTTTACCCCAAGGGTTCAAGGATATAGTCCCACTATTTTGGCCAGGCATTAGC	6922
Q	2945	CCAAAGCTTGAAGTCAATTTTTCATGCGGACACTTGTCTCTTGTGTAAGGAGTGAATTT	3004
D	6923	CCAAAGCTTGAAGTCTCAATTTTTCATGCGGACACTTGTCTCTTGTGTAAGGAGTGAATTT	6980
Q	3005	ACTTTTGCYGCYRTTCAAGAAACCTTGTGCACTCAAGCAACCCAGACRCTTTAAATTT	3064
D	6981	ACTTTTGCYGCYRTTCAAGAAACCTTGTGCACTCAAGCAACCCAGACRCTTTCAATTT	7040
Q	3065	CCTGCGVACTGTGGCTACAGAGTTTCCAAACABAGCTCABCTGTGCTCACAGCAGGT	3124
D	7041	CCTGCGVACTGTGGCTACAGAGTTTCCAAACABAGCTCABCTGTGCTCACAGCAGGT	7100

QY 3125 TAAATACCTTAGRCTAARATTAATCGAAGGCAACAGGGCCCTCAGTAGGAAYATCCA 3184
Db 7101 ----TACTTAGGGCTAAATTAATTCAAAGGCAACAGGGCCCTCAGTAGGAACACTCCA 7156
QY 3185 GCGCTACGCGGCTTATCCCTCATCYCAAAAACCTTAAAGCACTAAGGRRTTCTTGGCCT 3244
Db 7157 GCGTATACGCGCTTATCCCTCATCCAAAACCTTAAAGCACTAAGGGGATCTCTGGCGCT 7216
QY 3245 AAYAGYTTCTGCGCAAAATGATTCGCCAGGTMGGGCAAAATAGCCAGSYCATTAATATA 3304
Db 7217 AATAGTTTCTGCCCAAAATGATTCGCCAGGTMGGGCAAAATAGCCAGGTMATTAATA 7275
QY 3305 CASTAATTAAGGAACTCGAAGAGCCAAATCCCATTTAATAGATGAYATCGAAGYMR 3364
Db 7276 CACTAATTAAGGAACTCGAAGAGCCAAATCCCATTTAATAGATGAYATCGAAGYMR 7335
QY 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAGYCCCAAGTGTAAAGYTTGCC 3424
Db 7336 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAGYCCCAAGTGTAAAGYTTGCC 7381
QY 3425 AACGGGCAAGACTTTTSTTYATAATGCAAGAAACAAAGAAACAGAAATACCTCTGAGAGTCC 3484
Db 7382 AACGGGCAAGACTTTTCTTCAATATGACAG-AAAAACAGAAATACCTCTGAGAGTCC 7440
QY 3485 TTACACAGRTCCAGAGGAGYAGAGCTTGAACCYRTGCGRYACCTGASTAAGGAAATGATG 3544
Db 7441 TTACACAGATCCAGAGGAGTGAAGCTTGAACCTGAGCATACCTGASTAAGGAAATGATG 7500
QY 3545 TAGTGGCAAGAGGTTGRCYATTTGTTTAYGGGTAGTGTGGCAGTAGCACTYKATGAT 3604
Db 7501 TAGTGGCAAGAGGTTGACCTCATTTGTTTACGGGTAGTGTGGCAGTAGCACTYKATGAT 7560
QY 3605 CTGAAGCAGTTAAATATATACAGGAGAGAGATCTTACGTGTGACATCTCATAGKGA 3664
Db 7561 CTGAAGCAGTTAAATATATACAGGAGAGAGATCTTACGTGTGACATCTCATATATGGA 7620
QY 3665 AYAGCACTACTCACTGTAAAGAGACTGTGTGCTGTGACAAACGTTTACTTAATATTC 3724
Db 7621 ATGAGCACTACTCACTGTAAAGAGACTGTGTGCTGTGACAAACGTTTACTTAATATTC 7680
QY 3725 AGGCTCAATTAATCTTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCACTTTAAACCAG 3784
Db 7681 AGGCTCAATTAATCTTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCACTTTAAACCAG 7740
QY 3785 YGCATTTCTTCCAGAGAAAGAAAGAAAGATTAATTAATCTGCAACATTAATTTCTC 3844
Db 7741 CCACATTTCTTCCAGAGAAAGAAAGAAAGATTAATTAATCTGCAACATTAATTTCTC 7800
QY 3845 AAACCTATGCGCACTCGAGGGGACCTTGTAGAGTTCCYTTGACTGATCGYACCTTCAAC 3904
Db 7801 AAACCTATGCGCACTCGAGGGGACCTTGTAGAGTTCCYTTGACTGATCGYACCTTCAAC 7859
QY 3905 TTGTATACGTAGAGAAAGTTCTTTGTAGAAAGAAAGAAAGCTTGAAGAGYGGGTATGCACTG 3964
Db 7860 TTGTATACGTAGAGAAAGTTCTTTGTAGAAAGAAAGAAAGCTTGAAGAGYGGGTATGCACTG 7919
QY 3965 GTCAGAGTAATAGGAATATTTGAAAGTAATCCCTCACTCGAGAACTAGTCTTAAAGT 4024
Db 7920 GTCAGAGTAATAGGAATATTTGAAAGTAATCCCTCACTCGAGAACTAGTCTTAAAGT 7979
QY 4025 GCAGAACTAATAGCTCATATYKGGGCACTAGAAATTAAGAGAAAGAAAGGAAATATTA 4084
Db 7980 GCAGAACTAATAGCTCATATYKGGGCACTAGAAATTAAGAGAAAGAAAGGAAATATTA 8039
QY 4085 TTTACGACTCTAATATAGCTTATCTAGTCTTCATGCTCCATGMRCAATATTSABAGA 4144
Db 8040 TATAAGACTCTAATATAGCTTATCTAGTCTTCATGCTCCATGMRCAATATTSABAGA 8099
QY 4145 AGGGAATTCCTAATCTTCYAGAGAGAACCTATCAMAATCAGAGAACCTTATAGAAATTA 4204
Db 8100 AGGGAATTCCTAATCTTCYAGAGAGAACCTATCAMAATCAGAGAACCTTATAGAAATTA 8159

QY 4205 TTAATGCGMGTACAGAAACCTAABAGGTGGAAGTCTTACACTGCTGCGGGCTATCAGAAA 4264
Db 8160 TTAATGCGMGTACAGAAACCTAABAGGTGGAAGTCTTACACTGCGGGGTATCAGAAA 8219
QY 4265 GAAAGAAAGGAAATATSAAGRAAYGCCAAGCAKATATTTGAAGCAAAAGAGCTGCA 4324
Db 8220 GAAAGAAAGGAAATATSAAGRAAYGCCAAGCAKATATTTGAAGCAAAAGAGCTGCA 8279
QY 4325 AGGCAAGACCTCCATTAAGAAATGCTTATTAACCTTCCCTTATAGATAGGTAAATCCCTTC 4384
Db 8280 AGGCAAGACCTCCATTAAGAAATGCTTATTAACCAACCTTATAGATAGGTAAATCCCTTC 8339
QY 4385 CGGGAACCAAGCCCGCACTACTCAGAGAGAAACAGAAATGGGAACTCAGAGG-CAG 4443
Db 8340 CGGGAACCAAGCCCGCACTACTCAGAGAGAAACAGAAATGGGAACTCAGAGAGGAG 8399
QY 4444 TTTTCTCCCTCGGGAAGGTTAGCCCACTGAAGAGGAAATTAATCTTTGGCCGCAACTAT 4503
Db 8400 TTTTCTCCCTCGGGAAGGTTAGCCCACTGAAGAGGAAATTAATCTTTGGCCGCAACTAT 8459
QY 4504 CCAATGAAATTAATTAATAACCTTCATCAAACTTTCACTTAGGCATCGATAGACCCA 4563
Db 8460 CCAATGAAATTAATTAATAACCTTCATCAAACTTTCACTTAGGCATCGATAGACCCA 8519
QY 4564 TCATATGCGCAATTAATTAATTAATTAATGCAAGGCTTTCAAACTATCAAGCABATAT 4623
Db 8520 TCATATGCGCAATTAATTAATTAATTAATGCAAGGCTTTCAAACTATCAAGCABATAT 8579
QY 4624 CAGGCTCTGTAATGTCAGABAAATATATCCCTGCTATATGCGCAAGCTCTTCAAG 4683
Db 8580 CAGGCTCTGTAATGTCAGABAAATATATATCCCTGCTATATGCGCAAGCTCTTCAAG 8639
QY 4684 ABAACAAABAACAGGCTTATCCCTGABABAACCTGCAACCTGATTTTACCAACAGCCC 4743
Db 8640 ABAACAAABAACAGGCTTATCCCTGABABAACCTGCAACCTGATTTTACCAACAGCCC 8699
QY 4744 AAACCTCAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4803
Db 8700 AAACCTCAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8759
QY 4804 GCGTCCCTGTAAGGACAGAAAGGCGCAAGAGGTATTAAGGCACTAGTTCAATGAATA 4863
Db 8760 GCGTCCCTGTAAGGACAGAAAGGCGCAAGAGGTATTAAGGCACTAGTTCAATGAATA 8819
QY 4864 ATTCCAGATTCGCACTTCCCGAGGCTTACAGAGTACAAATAGCCCTGCTTCCAGGCC 4923
Db 8820 ATTCCAGATTCGCACTTCCCGAGGCTTACAGAGTACAAATAGCCCTGCTTCCAGGCC 8879
QY 4924 ACAATTAACCAAGGAGTATCCAGAGGCTTATAGTATACATATCACTTACATCAGGCCCTGA 4983
Db 8880 ACAATTAACCAAGGAGTATCCAGAGGCTTATAGTATACATATCACTTACATCAGGCCCTGA 8939
QY 4984 AGGCAAGTCTCTCAGGGAAGGTGAGAAATGAAGAAATCAACAAAGGCACTTAATAA 5043
Db 8940 AGGCAAGTCTCTCAGGGAAGGTGAGAAATGAAGAAATCAACAAAGGCACTTAATAA 8999
QY 5044 AAGCAAAACCAAGAAACCAACTCAGATGCGCTGTCTGTGCTTATAGCTTAAAGAA 5103
Db 9000 AAGCAAAACCAAGAAACCAACTCAGATGCGCTGTCTGTGCTTATAGCTTAAAGAA 9059
QY 5104 ATCTGCAATTTCCCAAAAGAGAGCTTACAGGCTTATAGAAATGCTGTATGGAAGGCC 5163
Db 9060 ATCTGCAATTTCCCAAAAGAGAGCTTACAGGCTTATAGAAATGCTGTATGGAAGGCC 9119
QY 5164 TTTCAATACCAATAGACTTGTGCTTGAACCAAGACAGCACTTATGTGAGACATCACT 5223
Db 9120 TTTCAATACCAATAGACTTGTGCTTGAACCAAGACAGCACTTATGTGAGACATCACT 9179
QY 5224 CTTTACCAATATCAAGAGTTCTTAAACCTTACAAAGAACTATCCTTGAAAGAGG 5283
Db 9180 CTTTACCAATATCAAGAGTTCTTAAACCTTACAAAGAACTATCCTTGAAAGAGG 9239
QY 5284 GAAAGAACTATTCACCCMGTGACATGATTAATGTAAGTCCCTCTCTAATTTCCC 5343

Db	9240	GAAGAAAGCAATTCACCCCTTGACAGTGGATATTAGCAAGTCCTCTCCCTCAATTC	9299
Oy	5344	CATCCCTAGATACATCTCTGGGAAGAACCCCAACCAGTCAATTTATATACCCCACTGCG	5403
Db	9300	CATCCCTAGATACATCTCTGGGAAGAACCCCAACCAGTCAATTTATATACCCCACTGCG	9359
Oy	5404	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTGGAGTCAAACTCTGAGTATG	5463
Db	9360	TTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTGGAGTCAAACTCTGAGTATG	9419
Oy	5464	CAAGGAACCTGAAATATCCAGAGACAAAGCTTAGCTATTCCTGTGAACCTTAGAGGAT	5523
Db	9420	CAAGGAACCTGAAATATCCAGAGACAAAGCTTAGCTATTCCTGTGAACCTTAGAGGAT	9479
Oy	5524	TGCGCTGCTCTTCAACAAACAACAGAGAAAGTAAGTAATATATATATCCCAATG	5583
Db	9480	TGCGCTGCTCTTCAACAAACAACAGAGAAAGTAAGTAATATATATATATATATATAT	9538
Oy	5584	GSCTCCCTTATCATATTTTTCTCTKATGTGTSTTTTAACTCTTCACTCTCACTGCA	5643
Db	9539	GCCTCCCTTATCATATTTTTCTCTTACTGTCTTTTAACTCTTCACTCTCACTGCA	9598
Oy	5644	CCCCCTCATGCTCCCTGTATGACAGTATGCTCCCTTACCMAGATTTCTATAGGAAG	5703
Db	9599	CCCCCTCATGCTCCCTGTATGACAGTATGCTCCCTTACCMAGATTTCTATAGGAAG	9658
Oy	5704	CAGGTCCTCCGAAATATTTGATGCCCAATGATAGAGTCTTTSYAGGAAACCCCAAC	5763
Db	9659	CAGGTCCTCCGAAATATTTGATGCCCAATGATAGAGTCTTTSYAGGAAACCCCAAC	9718
Oy	5764	TTCACTGCCCAACCCCATATGCCCCCGACATGCTATCATCTGCACTCTTTCATGAT	5823
Db	9719	TTCACTGCCCAACCCCATATGCCCCCGACATGCTATCATCTGCACTCTTTCATGAT	9778
Oy	5824	GCAAAATCAATTTATTTGACAGGAAATATGATTAATCTTAGTGTCTTGAAGACCTTGA	5883
Db	9779	GCAAAATCAATTTATTTGACAGGAAATATGATTAATCTTAGTGTCTTGAAGACCTTGA	9838
Oy	5884	GTCATGCTGTGTGACCTTATCTTCAACCAATGCTGATGCTATGAGGAGTGTGCA	5943
Db	9839	GTCATGCTGTGTGACCTTATCTTCAACCAATGCTGATGCTATGAGGAGTGTGCA	9898
Oy	5944	GATCAGGCAAGAAACATGTAAAAAGTAATCTCCCACTCACCGGGTATCATGTC	6003
Db	9899	GATCAGGCAAGAAACATGTAAAAAGTAATCTCCCACTCACCGGGTATCATGTC	9958
Oy	6004	ACTCTAGCCCCCTACAAAGACTAGATCTTCAAACTACATGAAACCTCTCGTACCT	6063
Db	9959	ACTCTAGCCCCCTACAAAGACTAGATCTTCAAACTACATGAAACCTCTCGTACCT	10018
Oy	6064	ACTGCGCTGTGAAGCTATTTAATACCAAGCTCATGAGGCTCCATGAGTCTGAGCCAA	6123
Db	10019	ACTGCGCTGTGAAGCTATTTAATACCAAGCTCATGAGGCTCCATGAGTCTGAGCCAA	10078
Oy	6124	AACCTTACTAAGTGTGATATGCTCCCTCTGAACTTCAAGCCATATGTTCAATCCCT	6183
Db	10079	AACCTTACTAAGTGTGATATGCTCCCTCTGAACTTCAAGCCATATGTTCAATCCCT	10138
Oy	6184	GTAACCTGAACATGGAACACTTACAGCAGAAATAAACACACTTCCGTTTATAGTGA	6243
Db	10139	GTAACCTGAACATGGAACACTTACAGCAGAAATAAACACACTTCCGTTTATAGTGA	10198
Oy	6244	CCTCTTGTTCCTCAATSTGGAATTAACCATACCTCAACCTCACTGTGTAAATTTAGC	6303
Db	10199	CCTCTTGTTCCTCAATSTGGAATTAACCATACCTCAACCTCACTGTGTAAATTTAGC	10258
Oy	6304	AATATCATATACAAACCACTCCCAATGATCAAGTGGTATCTCTCCCAACAATA	6363
Db	10259	AATATCATATACAAACCACTCCCAATGATCAAGTGGTATCTCTCCCAACAATA	10318
Oy	6364	GTCTGCTTACCTCAGGAATATTTTGTCTGTGATCTCAGCTTATCGTTTGAAT	6423

Db	10319	GTCTGCTACCCCTGAGAAATATTTTGTCTGTGTACCTGAGCTATCGTGTGTGAAT	10378
Qy	6424	GGCTCTTCAGAAATCTAATGTGCTTCTCTCAATTTCTTAAGCCCCCATATBRCAATCTACACT	6483
Db	10379	GGCTCTTCAGAAATCTAATGTGCTTCTCTCAATTTCTTAAGCCCCCATATBRCAATCTACACT	10438
Qy	6484	GAAACAAGTTTATACAGTTAATGTGCATATCTAAGCCCCGCAACAAAGAAGTACCCATTCTT	6543
Db	10439	GAAACAAGTTTATACAGTTAATGTGCATATCTAAGCCCCGCAACAAAGAAGTACCCATTCTT	10498
Qy	6544	CCTTTGTGTATAGAGCAGAGAGTGTCTAGTGTGCACTAGTACTGGCATTGGCCGTATCACA	6603
Db	10499	CCTTTGTGTATAGAGCAGAGAGTGTCTAGTGTGCACTAGTACTGGCATTGGCCGTATCACA	10558
Qy	6604	ACCTCTACTCAGTTTCTACTACAACTATCTCAAGAATTAATGGGGACATGGAAGGGTTC	6663
Db	10559	ACCTCTACTCAGTTTCTACTACAACTATCTCAAGAATTAATGGGGACATGGAAGGGTTC	10618
Qy	6664	GGCGAATCCCGGTGTCACTCTGCAAGATCAACTTAATCCCTAGCAGCAGTATGCTCTCTCA	6723
Db	10619	GGCGAATCCCGGTGTCACTCTGCAAGATCAACTTAATCCCTAGCAGCAGTATGCTCTCTCA	10678
Qy	6724	AATCGAAGAGCTTTTGAAGCTTGTCTAAGCTGARAAGGGGGAACTGTATTATTTTAAAGG	6783
Db	10679	AATCGAAGAGCTTTTGAAGCTTGTCTAAGCTGARAAGGGGGAACTGTATTATTTTAAAGG	10738
Qy	6784	GAAAGATGCTGTATTATATGTAAATCATCCGGAAATCGTCACTGAGAAAGTTTAAAGAAAT	6843
Db	10739	GAAAGATGCTGTATTATATGTAAATCATCCGGAAATCGTCACTGAGAAAGTTTAAAGAAAT	10798
Qy	6844	CSAAGATCGAATACAAACGTAKAGAGABAGGCTGAAAACATCGAACCCCTGGGGCTCCTC	6903
Db	10799	CSAAGATCGAATACAAACGTAKAGAGABAGGCTGAAAACATCGAACCCCTGGGGCTCCTC	10858
Qy	6904	AGCCATGAGATGACCCTGTGATTCCTCCCTCTTGAAGACCTCTAGACATAATATTGCTA	6963
Db	10859	AGCCATGAGATGACCCTGTGATTCCTCCCTCTTGAAGACCTCTAGACATAATATTGCTA	10918
Qy	6964	CTCCTCTTTGGAACCTGTATCTTTTACCTCTTGTTTAACTTTGTCTCTCCAGAAATCGAA	7023
Db	10919	CTCCTCTTTGGAACCTGTATCTTTTAACTCTTGTTTAACTTTGTCTCTCCAGAAATCGAA	10978
Qy	7024	GCTGTAAACATACAAATGAGGCCAAGATGACGTCAAGACTAAGATCTACCGCAGACCC	7083
Db	10979	GCTGTAAACATACAAATGAGGCCAAGATGACGTCAAGACTAAGATCTACCGCAGACCC	11038
Qy	7084	CTGAGCGGGCTGTGATGAGCCACAGATCTATGTTAATGACATCAAAAGGCCACCCCTCTGAG	7143
Db	11039	CTGAGCGGGCTGTGATGAGCCACAGATCTATGTTAATGACATCAAAAGGCCACCCCTCTGAG	11098
Qy	7144	GAAATCTCAGCTGACAACTCTCTACTACGCCCCCAATTGACAGGAAGCATTTAGACGGT	7203
Db	11099	GAAATCTCAGCTGACAACTCTCTACTACGCCCCCAATTGACAGGAAGCATTTAGACGGT	11158
Qy	7204	SGTGGGCCAACTCTCCCAACAGACACTTAGGTTTTCTGTGTGAGTGGGGGACTGAGAGAC	7263
Db	11159	SGTGGGCCAACTCTCCCAACAGACACTTAGGTTTTCTGTGTGAGTGGGGGACTGAGAGAC	11218
Qy	7264	AGGACTAGGTGATTTCCTAGGCTGATATGATGATCCCTTAAGCCTTAGCGGAAGGTGACC	7323
Db	11219	AGGACTAGGTGATTTCCTAGGCTGATATGATGATCCCTTAAGCCTTAGCGGAAGGTGACC	11278
Qy	7324	ACATCCACCTTTTAAACAGGGGGCTTGCAACTTAGTGTACACCTGACCAATCAGAGAGCTC	7383
Db	11279	ACATCCACCTTTTAAACAGGGGGCTTGCAACTTAGTGTACACCTGACCAATCAGAGAGCTC	11338
Qy	7384	ACTAAATATGCTAATTAAGGTAAAGACAGAGGTAAAGAAATAGCCAAATCATATTTGCTG	7443
Db	11339	ACTAAATATGCTAATTAAGGTAAAGACAGAGGTAAAGAAATAGCCAAATCATATTTGCTG	11398
Qy	7444	AAGAGCAGCAGAGAGAGGAAATATGATCGGGATTTTAAACCAAGTGTTCGAGCCGGCAACGG	7503
Db	11399	AAGAGCAGCAGAGAGAGGAAATATGATCGGGATTTTAAACCAAGTGTTCGAGCCGGCAACGG	11458

Db	3352	TGCTCGGGGCAAGGCGCATCCCAATGCAATCACTCCACAGAGCCCTGGGTATGGCTTAAC	3411
OY	2105	ATTGAGGGCCAGAAAGTGTCTCTCGGACACTGGTGGGTCTTTAGTCTTACTCTTC	2164
Db	3412	ATTAGAGGCCAGG-AGGTGTCTCTCGGACACTGGTGGGTCTTTAGTCTTACTCTTC	3470
OY	2165	TGTCCCGGACACTGTCTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGCGGCA	2224
Db	3471	TGTCCCGGACAACGTCTCTCCAGATCTGTCACTA-TCGAGGGGGTCC-TAAGACGGGCA	3528
OY	2225	GTCACTAAGTACTTTTCTTCCAGCCACTAAGTATGAACTGGGGAGCTTATCTTTTAC	2284
Db	3529	GTCACTAAGTAC-TTCTCCAGCCACTAAGTATG-ACTGGGGAGCTTATCTTTTAC	3586
OY	2285	ATGCTTTTCTAATTATGCTTGAAGAGCCCACTACCTTTGTTAGGAGAGACATTTAGCAA	2344
Db	3587	ATGCTTTTCTAATTATGCTTGAAGAGCCCACTACCTTTGTTAGGAGAGACATTTAGCAA	3646
OY	2345	AAGCAGGGGCAATTATACACTGAACATAGAGAAAGAACACCGTGTGTGTCCTG	2404
Db	3647	AAGCAGGGGCAATTATACACTGAACATAGAGAAAGAACACCGTGTGTGTCCTG	3705
OY	2405	CTTGAGAGAAAGATTAATCTTGAAGTCTGGGCAACAGAGACAAATAGACGAGCCAA	2464
Db	3706	CTTGAGAGAAAGATTAATCTTGAAGTCTGGGCAACAGAGACAAATAGACGAG-CAAA	3764
OY	2465	GAATGCCCGTCTGTTCAGGTTAACTAAAGGATTCACCTTCCTTCCCTACAAAGCA	2524
Db	3765	GAATGCCCGTCTGTTCAGGTTAACTAAAGGATTCACCTTCCTTCCCTACAAAGCA	3824
OY	2525	GTAACCCCTCAGACCCAGAGGCCCAACAGATTCGCAAAAGATGTTAAGACTTAAAGC	2584
Db	3825	GTAACCCCTCAGACCCAGAGGCCCAACAGACTCCAAAGATGTTAAGACTTAAAGC	3884
OY	2585	CCAAAGCTTAGTAAACCATGATTAATCTCCGAGTAATTCGTAAGTATGAGAGG	2644
Db	3885	CCAAAGCTTAGTAAACCATGATTAATCTCCGAGTAATTCGTAAGTATGAGAGG	3937
OY	2645	CACAGAAACCCAGTGGACAGTGGAGGTTAGTCAAGATCTCAGGATTAATCAATGAGGC	2704
Db	3938	TACAGAAACCCAGACAGACATGTA-GGTTAGTCAAGATCTCAGGATTAATCAAT-GAGGC	3995
OY	2705	CGTTGCTCTTTATATCCAGGCTGTACTAGCCCTTATATCTGTGATTTCCAAATACAGA	2764
Db	3996	TGTTGTTCTCTATAGCCAGCGTATACCTAGCCCTTATATCTGTGATTTCCAAATACAGA	4055
OY	2765	GGAAACAGATGCTTTTAACTCTGGAACCTTCAGAGATGCTTCTCATCTCCGTACA	4115
Db	4056	GGAAACAGATGCTTTTAACTCTGGAACCTTCAGAGATGCTTCTCATCTCCGTACA	4115
OY	2825	TCCGTACTCTCAATCTTGTGTTGCCCTTGAAGATATCTTCAAAACCCAGATCTCACTAC	2884
Db	4116	TCCGTACTCTCAATCTTGTGTTGCCCTTGAAGATATCTTCAAAACCCAGATCTCACTAC	4175
OY	2885	CTGACATTTTAAACCCAAAGGTTCAAGGATAGTCCCATCTAATTTGGCAGGACTTAGC	2944
Db	4176	CTGACATTTTAAACCCAAAGGTTCAAGGATAGTCCCATCTAATTTGGCAGGACTTAGC	4235
OY	2945	CCAAAGCTTAGTCAATCTTGTGGAACCTTGTGATCAAGCAACCAAGACCTTAAATTT	3004
Db	4236	CCAAAGCTTAGTCAATCTTGTGGAACCTTGTGATCAAGCAACCAAGACCTTAAATTT	4293
OY	3005	ACTTTTGGTGGCTTTCAGAAACCTTGGCATCAAGCAACCAAGACCTTAAATTT	3064
Db	4294	ACTTTTGGTGGCTTTCAGAAACCTTGGCATCAAGCAACCAAGACCTTAAATTT	4355
OY	3065	CTCGCACTGCTGGCTACAMGGTTCCAAACSAARAGCTCACTCTGCTCACAGAGT	3124
Db	4356	CTCGCACTGCTGGCTACAMGGTTCCAAACSAARAGCTCACTCTGCTCACAGAGT	4411
OY	3125	TAAATACTTAGRCCTAARATTAATCCAAAGCACARGGCCTCAGTGAAGAAATATCA	3186

Db	4414	----	TACTTAGGGCTAAATAATTATCCAAAGGACCCAGGCGCTCAGTAGGAAACATCC	4466
Oy	3185	GCCATATCTGGCTTATCTCATCTCYCAAAACCCCTAAAGCACTTAAGGRTCTCTGGC	3244	
Db	4470	GCCATATCTGGCTTATCTCATCTCAAAACCCCTAAAGCACTTAAGGAGATCTCTGGC	4529	
Oy	3445	AAVAGGTTCTGCGCAAMATGATTCCTCCAGTGTGGCBAATATAGCCAGGCTTATWATA	3304	
Db	4530	AATAGGTTCTGCGCAAMATGATTCCTCCAGTGTGGCBAATATAGCCAGGCTTATWATA	4588	
Oy	3305	CATTAATTAAGAAACTCAGAAAGCCAAATACCATTTATTAATGATGCAATCTGAAGTMR	3364	
Db	4589	CATTAATTAAGAAAGCTCAGAAAGCCAAATACCATTTATTAATGATGCAATCTGAAGTAG	4648	
Oy	3365	AAAGTGGCTTTCAGAGCCCTCTAAAGAAAGGCTTAAACCCAAAGTCCAGTGTATAGTGGC	3424	
Db	4649	AAAGTGGCTTTCAGAGCCCTCTAAAGAAAGGCTTAAACCCAAAGTCCAGTGTATAGTGGC	4694	
Oy	3425	AAAGGGGCAAGACTTTTSTTATATATATCAGAGAAAGAAACAGAAATVAGCTCTGAGTCC	3484	
Db	4695	AAAGGGGCAAGACTTTTCTCATATGTCACAG-AAAAACAGAAATAGCTCTTAGAGTCC	4753	
Oy	3485	TTACACAGATCCAGAGGAGYAGCTTGCAACCTATGGCRAVACTGAATAGAAATYAGT	3544	
Db	4754	TTACACAGATCCAGAGGAGTAGGCTTGCAACCTGAGGACATCTGACTTAAGAAATYAGT	4813	
Oy	3545	TAGTGGCAAAAGGTTGCTCATTTGTTATYGGGTATGGGTAGTGGCAAGTGGCACTTAT	3604	
Db	4814	TAGTGGCAAAAGGTTGCTCATTTGTTATYGGGTATGGGTAGTGGCAAGTGGCTTATAT	4873	
Oy	3605	CTGAAGCAGTTAAATAATATCAGGAGAGATCTTACTGTGGACATCTCATGAKTGA	3664	
Db	4874	CTGAAGCAGTTAAATAATATCAGGAGAGATCTTACTGTGGACATCTCATGATGTA	4933	
Oy	3665	AATGCACTACTCTGCTAAAGAGACTGTGGCTGTCAGACAACTGTTTACTTAATATC	3724	
Db	4934	AATGCACTACTCTGCTAAAGAGACTGTGGCTGTCAGACAACTGTTTACTTAATATC	4993	
Oy	3725	AGGCTCTATTACTTGAARGGCAGTGTCTGCACCTGTGCACCTTGAACCTTAACCG	3784	
Db	4994	AGGCTCTATTACTTGAARGGCAGTGTGTGCACCTGTGCACCTTGAACCTTAACCG	5053	
Oy	3785	YONCAATTCCTCCAGACATAGAAAGAAAGATATAATATCTGCAACAATATATCTC	3844	
Db	5054	COCAATTCCTCCAGACATAGAAAGAAAGATATAATATCTGCAACAATATATCTC	5113	
Oy	3845	AAACCTATGCACTCGAGGGGACCTTATAGAGTTCCTTGAATCTGATCTGACCTTCAAC	3904	
Db	5114	AAACCTATGCACTCGAGGGGACCTTATAGAGTTCCTTGAATCTGATCTGACCTTCAAC	5172	
Oy	3905	TTGTATACTGATGAGAGTTCCTTGTATGAAAGAAAGAGACTTGCAGAAAGTGGGTATG	3964	
Db	5173	TTGTATACTGATGAGAGTTCCTTGTATGAAAGAAAGAGACTTGCAGAAAGTGGGTATG	5232	
Oy	3965	GTCAGTGTATATGAAATATTTGAAAGATCCCTCACTCCAGAGAACTATAGTCTAGCTR	4024	
Db	5233	GTCAGTGTATATGAAATATTTGAAAGATCCCTCACTCCAGAGAACTATAGTCTAGCTA	5292	
Oy	4025	GCAGAACTAATAGCCCTTCATYKGGGCACTAGAAATTAGAGAGAAAAAGGTYAATATA	4084	
Db	5293	GCAGAACTAATAGCCCTTCATYKGGGCACTAGAAATTAGAGAGAAAAAGGTYAATATA	5352	
Oy	4085	TATACAGACTCTATATATGCTTACCTAGTCTGATCCCTATGMRGCAATATGABAGAA	4144	
Db	5353	TATACAGACTCTAATATATGCTTACCTAGTCTGATCCCTAGGCAATATGABAGAA	5412	
Oy	4145	AGGAAATTCCTAATCTTCGAGAGAAACCTATCMAACATCAGAAAGCATTAAGABATTA	4204	
Db	5413	AGGAAATTCCTAATCTTCGAGAGAAACCTATCMAACATCAGAAAGCATTAAGABATTA	5472	
Oy	4205	TTATYGGCGMTACAGAAACTATAGAGGTTGAGTCTTACACTGTGCGGGTCATCANAAA	4264	
Db	5473	TTATYGGCGMTACAGAAACTATAGAGGTTGAGTCTTACACTGTGCGGGTCATCANAAA	5532	

4265 GGAAGRAAAGGAAATASAGRAAYTGCACAGACATATTTGAGCAAAAAGAGCTGCA 4324
4325 AGGCAGGACCCCTCATTTAGAAATGCTTTATTAACTTCCCTTATAGATAGGGTAAATCCCTTC 4384
5593 AGGCAGGACCCCTCATTTAGAAATGCTTTATTAACTTCCCTTATAGATAGGGTAAATCCCTTC 5652
4385 CGGGAAACCAAGCCCCGAGTACTCAGCAGAGAAAACAAATGGGGAACCTCAGCAGG- CAG 4443
5653 CGGGAAACCAAGCCCCGAGTACTCAGCAGAGAAAACAAATGGGGAACCTCAGCAGG- CAG 5712
4444 TTTTCTCCCTCGGGAAGTGAAGCCTGAGCACTGAGAAAGGAAATATCTTTGCTGCACTAT 4503
5713 TTTTCTCCCTCGGGAAGTGAAGCCTGAGCACTGAGAAAGGAAATATCTTTGCTGCACTAT 5772
4504 CCAATGGAATTTACTTAAACCTTTCATCAACCTTTCACTTAGCGATGATAGCACC 4563
5773 CCAATGGAATTTACTTAAACCTTTCATCAACCTTTCACTTAGCGATGATAGCACC 5832
4564 TCARATGGCCAAATCATATTTATCTGACAGGCTTTTCAAAAATACTCAAGABATAT 4623
5833 TCAGATGGCCAAATCATATTTATCTGACAGGCTTTTCAAAAATACTCAAGABATAT 5892
4624 CAGGCTGTGAATGTGACABAAATTAATCCCTGCTATGCGCAAGCTCCTTCAAG 4683
5893 CAGGCTGTGAATGTGACABAAATTAATCCCTGCTATGCGCAAGCTCCTTCAAG 5952
4684 ARPAACAARPAACAGGCTTATACCTGTAPARAPACTGGCAATCTGATTTTACCAACAGCC 4743
5953 AGAACAARPAACAGGCTTATACCTGTAPARAPACTGGCAATCTGATTTTACCAACAGCC 6012
4744 AAACTCAGGATTTCTAGTATCTACTGCTGCTGATATCTTTCAAGGGTGGGCAAG 4803
6013 AAACTCAGGATTTCTAGTATCTACTGCTGCTGATATCTTTCAAGGGTGGGCAAG 6072
4804 GCTTCCCTGTAGAGCAGAAAAAGGCCAAGGTAATAAAGGCACTGTTTCATGAAATA 4863
6073 GCTTCCCTGTAGAGCAGAAAAAGGCCAAGGTAATAAAGGCACTGTTTCATGAAATA 6132
4864 ATTCCAGATTTGAGACTTCCCGAGGCTTACAGATGACAAATAGCCTGCTTTCAGGCC 4923
6133 ATTCCAGATTTGAGACTTCCCGAGGCTTACAGATGACAAATAGCCTGCTTTCAGGCC 6192
4924 ACAGTAAACCAAGGAGTATCCAGGCTTATAGTATAGATATCACTTACACTGGGCTG 4983
6193 ACAGTAAACCAAGGAGTATCCAGGCTTATAGTATAGATATCACTTACACTGGGCTG 6252
4984 AGGCACAGTCTCAGGGAAGTGCAGAAAAATGATGAAAVACTCAAGGACATCTTAA 5043
6253 AGGCACAGTCTCAGGGAAGTGCAGAAAAATGATGAAAVACTCAAGGACATCTTAA 6312
5044 AAGCAAAACCAAGGAAACCACTTCAATGCTGCTGCTGCTTATAGCTTAA 5103
6313 AAGCAAAACCAAGGAAACCACTTCAATGCTGCTGCTGCTTATAGCTTAA 6372
5104 ATCTGCAATTTTCCCAAAAAGCAGACTTATAGCCATACGAAAGCTGATAGAGGCC 5163
6373 ATCTGCAATTTTCCCAAAAAGCAGACTTATAGCCATACGAAAGCTGATAGAGGCC 6432
5164 TTCTATACCAATGACTTGTGCTGACCAAGACAGCACTTATGTCAGACATCCT 5223
6433 TTCTATACCAATGACTTGTGCTGACCAAGACAGCACTTATGTCAGACATCCT 6492
5224 CTTTACCAATGACTTGTGCTGACCAAGACAGCACTTATGTCAGACATCCT 5283
6493 CTTTACCAATGACTTGTGCTGACCAAGACAGCACTTATGTCAGACATCCT 6552
5284 GAAAAGAACTATTCACCCMGTGACATGATTAATGTCAGTCCCTTCCTTATATCC 5343
6553 GAAAAGAACTATTCACCCMGTGACATGATTAATGTCAGTCCCTTCCTTATATCC 6612

5344 CATCCCTAGATACATCTGGGAAGACCTTACCAGTCAATTTATATACCCCACTGCG 5403
6613 CATCCCTAGATACATCTGGGAAGACCTTACCAGTCAATTTATATACCCCACTGCG 6672
5404 TTAAATGCTGAGAGTGTGCTTGGATATCATCACTTTGATCAAAATCTGATATC 5463
6673 TTAAATGCTGAGAGTGTGCTTGGATATCATCACTTTGATCAAAATCTGATATC 6732
5464 CAAAGAACCTGAAAAATCCAGAGCAAGCTAGCTATCTGTGAACCTCTAGAGATT 5523
6733 CAAAGAACCTGAAAAATCCAGAGCAAGCTAGCTATCTGTGAACCTCTAGAGATT 6792
5524 TGGCTGCTCTTCAACAACAACAAGAGAAATTAATCAATAATCCCATG 5583
6793 TGGCTGCTCTTCAACAACAACAAGAGAAATTAATCAATAATCCCATG 6851
5584 GGCCTCCCTTATCATATTTTCTCTATGTTTSTTTCACCTTTTCACTCAGCA 5643
6852 GGCCTCCCTTATCATATTTTCTCTATGTTTSTTTCACCTTTTCACTCAGCA 6911
5644 CCCCCTCATGCGCTGTATGACCAAGTACTCCCTTACCMAGATTCTATGAGAAATG 5703
6912 CCCCCTCATGCGCTGTATGACCAAGTACTCCCTTACCMAGATTCTATGAGAAATG 6971
5704 CAGGCTCCGGAATATTTGATGCCCCATGATATGAGAGTCTTTTAAAGGAACCCCA 5763
6972 CAGGCTCCGGAATATTTGATGCCCCATGATATGAGAGTCTTTTAAAGGAACCCCA 7031
5764 TTCACTGCGCAACCCATATGCCCCGCACTGCTATCACTGCGCACTTTGATGAT 5823
7032 TTCACTGCGCAACCCATATGCCCCGCACTGCTATCACTGCGCACTTTGATGAT 7091
5824 GCATAATCATATATGAGCAGAAAAATGATTAATCTGATGTTGCTGAGAGACTTGA 5883
7092 GCATAATCATATATGAGCAGAAAAATGATTAATCTGATGTTGCTGAGAGACTTGA 7151
5884 GTCACTGCTGTTGGACTTACTTACCCAACTGCTATGCTGATGGGGTGAAGTTCAA 5943
7152 GTCACTGCTGTTGGACTTACTTACCCAACTGCTATGCTGATGGGGTGAAGTTCAA 7211
5944 GATCAGGCAAGAAAAATGATTAATGATATCCCACTCACGSGGATACATGAG 6003
7212 GATCAGGCAAGAAAAATGATTAATGATATCCCACTCACGSGGATACATGAG 7271
6004 ACCTCAGGCTTACCAAGAGTATCTCTCAAAAATACTATGAAACCTCTGATCCAT 6063
7272 ACCTCAGGCTTACCAAGAGTATCTCTCAAAAATACTATGAAACCTCTGATCCAT 7331
6064 ACTGCTGCTGTAAGCTTATTAATACACCTCTCACTGGGCTCATGAGGCTCGGCCAA 6123
7332 ACTGCTGCTGTAAGCTTATTAATACACCTCTCACTGGGCTCATGAGGCTCGGCCAA 7391
6124 AACCTCAGTATGCTGATGATGCTCCCTCGAATCTTACGAGGATGTTCAATCCCT 6183
7392 AACCTCAGTATGCTGATGATGCTCCCTCGAATCTTACGAGGATGTTCAATCCCT 7451
6184 GTACCTGAAACAATGGAACAATCTTACAGCAGAAATTAACAACAATCCGTTTATGAGA 6243
7452 GTACCTGAAACAATGGAACAATCTTACAGCAGAAATTAACAACAATCCGTTTATGAGA 7511
6244 CTTCTGTTTCCATSTGGAATTAACCCATACCTCAACCTCACTGCTGTAATTTAGC 6303
7512 CTTCTGTTTCCATSTGGAATTAACCCATACCTCAACCTCACTGCTGTAATTTAGC 7571
6304 AATACACATACACACCAACCACTCCCAATGATGAGGTTAACTGCTCCACACAAATA 6363
7572 AATACACATACACACCAACCACTCCCAATGATGAGGTTAACTGCTCCACACAAATA 7631
6364 GTCTGCTACCTCAGAAATATTTTGTCTGTGATACCTCAGCCTATGCTTGTGAAT 6423
7632 GTCTGCTACCTCAGAAATATTTTGTCTGTGATACCTCAGCCTATGCTTGTGAAT 7691
6424 GGCTTTCAAAATCTATGTGCTTCTCTCATTTCTATGAGGCCCTTATGCACT 6483

Db	7692	GGCTCTTGAAATCTATGTCGCTTCCCTCACTTCTTAGTGCCCTATGACATCTACACT	7751
Qy	6484	GAAACAAGATTATACAGTTATGTCATATCTTAAGCCCCCGCAACAAAGATACCCATTCTT	6543
Db	7752	GAAACAAGATTATACAGTTATGTCATATCTTAAGCCCCCGCAACAAAGATACCCATTCTT	7811
Qy	6544	CCTTTTGTTATAGGACAGAGAGTCATAGTGCACTAGTACCTGGCAATTGGCCGTATCACA	6603
Db	7812	CTTTTGTATATAGGACAGAGAGTCATAGTGCACTAGTACCTGGCAATTGGCCGTATCACA	7871
Qy	6604	ACCTCTACTCAGTTCTACTACAAACATCTCAAGAACTAAATGGGGACATGAAAGGGATC	6663
Db	7872	ACCTCTACTCAGTTCTACTACAAACATCTCAAGAACTAAATGGGGACATGAAAGGGATC	7931
Qy	6664	GCCGACTCCCTGGTCACTTGCAAGATCAACTTAAGTCCCTAGCAGAGTATGCTTGTA	6723
Db	7932	GCCGACTCCCTGGTCACTTGCAAGATCAACTTAAGTCCCTAGCAGAGTATGCTTGTA	7991
Qy	6724	AATCGAAGAGCTTTAGACTTGCTAACCCGCTGAPAGAGGGGGAACTGTTATTTTATAGGG	6783
Db	7992	AATCGAAGAGCTTTAGACTTGCTAACCCGCTGAAAGAGGGGGAACTGTTATTTTATAGGG	8051
Qy	6784	GAAAGATGCTGTTATTAATGTTATCAATCCGGAATGCTCACTGAGAAAGTTAAGAATTT	6843
Db	8052	GAAAGATGCTGTTATTAATGTTATCAATCCGGAATGCTCACTGAGAAAGTTAAGAATTT	8111
Qy	6844	CSAGATCGAATCAACAGTAGAGAGAGACTCGAAAACATCGAACCCCTGGGCGCTCTC	6903
Db	8112	CGAGATCGAATCAACAGTAGAGAGAGACTCGAAAACATCGAACCCCTGGGCGCTCTC	8171
Qy	6904	AGCCBATGGATGCCCTGGATTCCTCCCTTTTAGACCTCTAGACGTAATATTTGCTA	6963
Db	8172	AGCCBATGGATGCCCTGGATTCCTCCCTTTTAGACCTCTAGACGTAATATTTGCTA	8231
Qy	6964	CTCCTCTTTGGACCCCTGTAATCTTTTAACTCTCTGTTTAACTTTGTCTCTTCCAAATCGAA	7023
Db	8232	CTCCTCTTTGGACCCCTGTAATCTTTTAACTCTCTGTTTAACTTTGTCTCTTCCAAATCGAA	8291
Qy	7024	GCTGTAAACTCAAAATGAGGCCCAAGATGCACTCCCAAGCTAAGATCTACCGCAAGCC	7083
Db	8292	GCTGTAAACTCAAAATGAGGCCCAAGATGCACTCCCAAGCTAAGATCTACCGCAAGCC	8351
Qy	7084	CTGCAACGGGCTGTATGAGCCACGATCTGATGTTATATGACATCAAGAGCAACCCCTCTGAG	7143
Db	8352	CTGCAACGGGCTGTATGAGCCACGATCTGATGTTATATGACATCAAGAGCAACCCCTCTGAG	8411
Qy	7144	GAAATCTCAGCTGACAACTCTTAATGCGCCCAATTCAGCAGAGACAGTATGAGCGGT	7203
Db	8412	GAAATCTCAGCTGACAACTCTTAATGCGCCCAATTCAGCAGAGACAGTATGAGCGGT	8471
Qy	7204	SGTCGGCCAACTCCCCCAACAGACATTAAGTTTTCCTGTTGAGATGGGGGAC	7255
Db	8472	SGTCGGCCAACTCCCCCAACAGACATTAAGTTTTCCTGTTGAGATGGGGGAC	8523

```

RESULT 3
US-09-949-016-17417/C
Sequence 17417, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08

```

```

? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 17417
? LENGTH: 77772
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)---(77772)
? OTHER INFORMATION: n = A,T,C or G
? US-09-949-016-17417

```

Query Match	39.5%	Score	29.94.2	DB	4	Length	77772	
Best Local Similarity	85.5%	Pred.	No. 0					
Matches	3410	Conservative	149	Mismatches	362	Indels	69	Gaps
QY	1812	ACCTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTGAATGCTCTPATAGGCTTGCT	1871					
Db	19931	ACCTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTGAATGCTCTPATAGGCTTGCT	19878					
QY	1872	TCCAGTGCAGTCTACAAAGGACATTTTAAAAAAGTTGTCCAGTAAAGTAAGCGCGCCC	19311					
Db	19871	TCCAGTGCAGTCTACAAAGGACATTTTAAAAAAGTTGTCTGAATTAAGTAATTAAGTCGCC	19811					
QY	1932	TTGGTCGATGCCCCCTTATTTTCAAGGGAACTACTGAAAGGCCACTGGCCCCAGGGGACAAA	19911					
Db	19811	-TATTCATGCCCCCTTATGTCTAAAGGGAACTACTGAAAGGCCACTGGCCCCAGGGGATGAA	19751					
QY	1992	GGCTTTTGAATGAGAGAGCCACTTAACCAATGATTCAGCAGCAGAGACTGAGGGTGCTGG	2051					
Db	19752	GGTCTCTTGAATGAGAGAGCCACTTAACCAATGATTCAGCAGCAGAGACTGAGGGTGCTGG	19666					
QY	2052	GGCAAGCGGCATATCCCATGCAATCACTCACAGAGCCCTGGGTATGCTTGAACAATTAGG	2111					
Db	19692	GGCAAGCGGCATATCCCATGCAATCACTCACAGAGCCCTGGGTATGCTTGAACAATTAGG	19666					
QY	2112	GCCAGGAAGG---TGTCTCTGGACACTGGTGCGGCTTCTTATGCTTACTCTCTCTC	21666					
Db	19649	GCCAGGAAGGTTAACTGTCTCTCTGAACCTGGCATGCGCTTCTCATGCTTACTCTCTGCT	19551					
QY	2169	CCGGAACAACGTGCTCCAGATCTGTCACTATTCTGAAGGGGGTCCNTAAGACGGGCACTCA	2222					
Db	19589	CTGGAACAACGTGCTCCAGATCTGTCACTATTCTGAAGGGGGTCCNTAAGACGGGCACTCA	19531					
QY	2229	CTAGATACCTTTTCCCAAGCACTPATGTTATGAATCGGGAGCTTATTTCTTTTCAATGC	2288					
Db	19532	CTAGATACCTTTTCCCAAGCACTPATGTTATGAATCGGGAGCTTATTTCTTTTCAATGC	19471					
QY	2289	TTTTCTAATTAATGCTTGAAGCCCCCACTACCTTGTTAGGGAGACACTTCAGCAAAAGC	2344					
Db	19474	CTTTCTAATTAATGCTTGAAGCCCCCACTTCCTTGAAGGATACATTTTGAACAAAAGC	19414					
QY	2349	AGGGGCAATTAATACCTGAAACATAGGAGAAAGAACCCGTTTGTGTGTCCTGCTTG	2406					
Db	19414	AGGGGCAATTAATACCTGAAACATAGGAGAAAGAACCCGTTTGTGTGTCCTGCTTG	19331					
QY	2409	AGGAAGGAATTAATCTTGAAGTCTGGGCAACAGAGACAAATATATGAGCAGGCCAAAGAT	2466					
Db	19356	AGGAAGGAATTAATCTTGAAGTCTGGGCAACAGAGACAAATATATGAGTGAAG-AGAAAGAT	19221					
QY	2469	GCCGCTCTCTGTCAAGTTAACTPAAAGATTCCACTTCCCTCCCTACCAAGGCAAGTAC	2522					
Db	19297	GCCGCTCTCTGTCAAGTTAACTPAAAGATTCCGCTCTCTTCCCAACAAAGGCAAGTAC	19221					
QY	2529	CCCTCTGAGACCCAGGCGCCAAACAAGATTCCAAAGAATTGTTAAGACTTAAAAAGCCCA	2588					
Db	19237	CTTCTTGAACCCAGGCGCCAAACAAGACTCCAAAGAATTGTTAAGACTTAAAAAGCCCA	19111					
QY	2589	GGCTTATTAATAACCATGCAATCACTCCCTGCAAGTATTCCTGATGAGATTGAAGAGCA	2644					
Db	19177	GGCTTATTAATAACCATGCAATCACTCCCTGCAATCACTCC-----AATTTTGAAGTAC	19111					
QY	2649	GAAACCAAGTGAACATGGAAGGTTATGTCAGATCTTCAGATTATCAATGGAAGCGGTT	2706					

Db	19124	GAATCCAAATGGACAGTGG- GGTTAGTGTGCAAGATCTCAGGATTTCAAT- GAGGCTGTT	190677
Qy	2709	GTCCCTTTTAACCCAGCTGTACCTAGCCCT--TAACTGTGMYTTCGCAATACAGAG	2765
Db	19066	GTTCCTCTAATACCCAGCTGTACTAATCTTAACTACACTGTGCTTTCGCAATACAGAG	190077
Qy	2766	GAAGCAGAGTGTGTACASTCTGTGACCTTAAAGATGCGCTTTCTTGACCTCTGTACAT	2825
Db	19006	GAAACAGAGTGTGTACAGTCTGTGACCTTAAAGATGCGCTTTTAACTGACCTCTGTACAT	189477
Qy	2826	CCGACCTCAATCTTGTTGCGCTTGAAGTACTTGAACCCGACATCTCACTCAC	2885
Db	18946	CTGACTCTCAATCTTGCTTGCTGCTTTGAAGATCTTGAATCTAATGTCTCAACACACC	188877
Qy	2886	TGGACTRTTATACCCCAAGGTTGAGGATAGYCCCATCTAATTGAGCAGGACATTAGCC	2945
Db	18886	TGGACTGTTTACCCCAAGGTTTATGACAGGCCCATCTAATTGAGCAGGACATTAGCC	188277
Qy	2946	CAAGACTTGAAGYCARVTMTACATACCTGTGACACTCTGTCTCTTGATGTGATATTTA	3005
Db	18826	CGAGACTTGAAGCAAGTCTCAATCTTGGGCACTCTTGTCTCTTGATGTGATATTTA	187677
Qy	3006	CTTTTGGYCCYCTRTTGAAGAACTTGTCATAGGACCAACGACCTCTTMAATTTC	3065
Db	18766	CTTTTAGCCACCCATTCAGAAACCTTGCCATCAGCACCCCAAGTCTCTTAACTTC	187077
Qy	3066	CTCGCYACCTGTGGCTACAGMGTTTCCAAASABARGCTCARCTGCTCAGCAGAGTT	3125
Db	18706	ATGGCCACCTGTGGCTACAGAGTTTCCAGACCAAGGCTCAGCTCTGCTCAGCAGAGTT	186477
Qy	3126	AAATACTTAGGCTAATATATCCAAAGGCACCAAGGCTCTCAGTAGAGAAVYATCAG	3185
Db	18646	AAATACTTGAAGGCTAATATATCCAGACACACCAAGGCTCTCAGTAGAGAAVYATCAG	185877
Qy	3186	CCATATCTGGCTATCCTCATVCMAAACCTTAAAGCACTAAGGRTTCTCGGCKTA	3245
Db	18586	CCATATCTGGCTATCCTTATCCCAAAACATTAAGCAATTAAGGRTTCTCGGCKTA	185277
Qy	3246	AYAGYTTTCGCGCAAAWTGATATCCCCAGGTTWTGGCAAAATAGCAGAGYCATTMATAC	3305
Db	18526	ACAGGCTTCGCGCAAAWTGATATCCCCAGGTTWTGCAAAATAGCAGGCTATTAAC	184687
Qy	3306	ASTAATTTAAAGAACTCAGAAAGCCAAATCCCATTTAATTAAGATGAGVAMCTGAAGYMA	3365
Db	18467	ACTAATTTAAAGAACTCAGAAAGTGAATACCCATTTAGTAAGAAAGACACCTGAAGCAGA	184087
Qy	3366	AGTGGCTTTCAGAGCCCTTAAGAAAGGCTTAAACCCAGTCCAGGTGTAAAGTTGGCA	3425
Db	18407	AGTGGCTTTCAGAGCCCTTAAGAAAGGCTT--AATCCAAAGCCATGTTTAAAGTTGGCA	183507
Qy	3426	ACRGGGCAAGACTTTTSTTATATATATACAGAAAAAAACAGAAVAGCTCTRGAGTCT	3485
Db	18349	AAGGGCAAGACTTTTCTTATATGTGTACAG-AAAAAACAGAAATAGCTCTTAGAATTCCT	182917
Qy	3486	TACACAGRTCCRAAGGAYGAGCTTGCAACCYRTGCGRYACCTGASTAGAGAAVYTAGGT	3545
Db	18290	TACACAGRTCCGAGGAGCCAGGTTTGCAACCCATGSCATCTGATGAAGAACTGATGT	182317
Qy	3546	AGTGGCAAGGTTGRCYTCATTTGTTAAYGGGTAGTGTGGCAAGTACAGATYKTAGTTC	3605
Db	18230	AGTGGCAAGGTTGRCCTCATTTGTATGTGGTATAGCAGACATAGCAGTCTTAGTATC	181717
Qy	3606	TGAAGCAGTTAAATATAACAGGGBAGAGATCTTACTGTGTGGACATCTCATGAKGTAA	3665
Db	18170	TGAAGCAGTTAAATATGATACAGGGBAGAGATCTTACTGTGTGTGATGTCTCATGATGTGA	181117
Qy	3666	YRGCACTACTACTGTAAAGAGACTTGTGGCTGTACAGACAACYGTTTACTTAATATATCA	3725
Db	18110	CGGCAATCTCACTGTAAAGAGACTTAAGGCTGTACAGACAACCGTTTCTTAATATATCA	180517
Qy	3726	GGCTCTATTACTTGAABGGCCAGTGTCTCPACTGTGCACTTGTGCAACTTTAAACCAAGY	3785

Db	18050	GGCTATTATTCTTGAAGGCCAGTCTGCTGCACTGCGGCACTGTCGAACTCTTAAATTCACG	17991
Qy	3786	CNCATTTCTTCAGACAATGAAAGAAAGTAAATATACGTCAACAACTAATATTTCTCA	3845
Db	17990	CACATTTCTTCAGACAATGAAAGAAAGTAAATATACGTCAACAACTAATATTTCTCA	17931
Qy	3846	AACCTATGCACTCGAGGGGACCTTATGAGTTCCTTTGACTGATCCYGACCTTCAACT	3905
Db	17930	AACCTATGCACTCGAGGGGACCTTATGAGTTCCTTTGACTGATCCYGACCTTCAACT	17872
Qy	3906	TGTATACTGATGGAAAGTTCTTTGTAGAAAAAGGCTTCGAAAATGCGGGTATGCACTGG	3965
Db	17871	TGTATACTGAGGAAAGTTCTTTGTAGAAAAAGGCTTCGAAAATGCGGGTATGCACTGG	17812
Qy	3966	TCAGTGAATATGSAATATATTTGAAAGTAATCCCTCACTCCAGAACTAGTGGTATGCTRG	4025
Db	17811	TCAGTGAATATGSAATATTTGAAAGTAATCCCTCACTCCAGAACTAGTGGTATGCTRG	17752
Qy	4026	CAGAACTAATAGCCCTCAATYKGGGCACTAGAAATTTAGAGAGRAAAAAAGGTAAATAT	4085
Db	17751	CAGAACTAATAGCCCTCACTAGGCACTAAATTTAGGAAAGGAAAAAGGTAAATATAT	17692
Qy	4086	ATACAGACTCRARATATCTYACCGTACGTCATGCCCATGMRCCAAATATSSAAGAAA	4145
Db	17651	ACACAGACTCTRAAGATGCTTAACCTAGTCTTCATGCCCCATGRCACCAATATGAGAGAAA	17632
Qy	4146	GGGAATTCCTAATCTTCYAGRGAACACCTATCMACATCAGGAAGCCATTAGABATTAT	4205
Db	17631	GGGAATTCCTAATCTTCGAGGGAAACACCTATCAAACTCAGGAAGCCATTAGAGATTAC	17572
Qy	4206	TATYGGCKGTACAGAAACCTTAPAGAGGTGMAAGTTTACACTGCTGGGCTCATMAAAG	4265
Db	17571	AATTGGCTGTACAGAAACCTTAAGGGGTGGCAGTCTTACACTGCGGGGTGTACAAAG	17512
Qy	4266	GAAAGRAAAGGAAATATASARGAATYGCCAAGCAKATATTGAACMAAAAGAGTCGAA	4325
Db	17511	GAAAGRAAAGGAAATATAGAGGAAACGCGCAAGACGATATCTGAAACCAAAAGAGCCGCA	17452
Qy	4326	GGCAGAGCCCTCCATAGATATGCTTATTAACCTCCCTAGTATAGGGTATTCCTTCC	4385
Db	17451	GGTGGAGCCCTCCATAGAAATGCTTATAGAAAGACCCCTAGTATGGGGTATCCCTGT	17392
Qy	4386	GGGAAACCAAGCCCCAGTACTCAGCAGAGAGAAACAAATGGGGAACCTCACAGAGCATAGTT	4445
Db	17391	AGGAAACCAAGCCCCAGTACTCAAAAGAAATAGGTGGGGAACCTCACAGAGCATAGTT	17332
Qy	4446	TTCTCCCTCGGAGCGTTAGCCACTGAGAAAGGGAATATCTTTGGCTGCAATATCC	4505
Db	17331	TTCTCCCTCGAGATGCTATGACCCACCGAAGAAAGAAATATCTTTGGCTGCAATATCC	17272
Qy	4506	AATGAAATTTCTTAAAAACCTTATCAACCTTTCACTTAGCGCATATAGCACCATC	4565
Db	17271	AATGAAATTTCTTAAAAACCTTATCAAGACTTTCATAGACTTTCATAGGCAATTGATAGCACTATC	17212
Qy	4566	ARATGGCCAAATCATTAATTTATCTGACCAAGGCTTTTCAAAACTATCAAGCARATATKTC	4625
Db	17211	AGAGGGCCAAATCATTTATTACTGATCAAGACTTTTCAAAACTATCAAGCAGATATGTC	17152
Qy	4626	GGGCTGTGAATGTGCGCARAAATTAATCCCTGCGCTTAATCCGCAACACTCCTTAGAG	4685
Db	17151	GGGCTGTGAATGTGCGCAAAATTAATCCCTGCGCTTATTCGCGCAACACTCCTTAGAG	17092
Qy	4686	AACAAAAAAGCGGCACTTACCTGTARARARACTGGCAACT-GATTATACCAAGGCCA	4744
Db	17091	AACAAAAAAGCGGCACTTACCGCAGAGAAAGCTGGCAACTGATTTTATCCCAATGCCCA	17032
Qy	4745	AACCTCAGGATTTCAAGTACTACTAGTCTGGGTARATATCTTTCACGGCTTGGGCARAG	4804
Db	17031	AATCTCAGGATTTCAATATCTACTAGTCTGGATAGATATCTTTCACTGGTGGGAGGAG	16972
Qy	4805	CCTTCCCTGTAGAGCAAGAAAAGGCCCAAGAGGTATTAAGGCACTAGTTTATGAAATTA	4864
Db	16971	CCTTTCCTTGTAGAGCAAGAAAAGGCCCAAGAGCTGAATTAAGGCACTAATTTATGAAATTA	16912

Oy	4665	TTCCAGATTGGACCTTCCCGAGGCTTAAAGAGTGCATATACCTCGCTTTCAGAGCCA	4924
Db	16911	TTCCAGATTGGACCTTCCCGAGGCTTAAAGAGTGCATATACCTCGCTTTCAGAGCCA	16852
Oy	4925	CAGTACCAGGAGTATCCAGCG---TTAGATATACGATATACATTACCTGCGCC	4980
Db	16851	CAGTACCAGGAGTATCCAGAGCGTTACTTATGAGGCAATACATATCACTTACCTGCGCC	16792
Oy	4981	TGAAGGCCACAGTCCCTACAGGAAAGGTGAGAAAAATGAAATGAAATCTCAAGACATCTA	5040
Db	16791	TGAGAGCCACATCTCTACAGGAAAGGTGAGAAAAATGAAACAAACCTCAATCTA	16732
Oy	5041	AAAAAGCAAAACCAAGAAAAACCACTCACATGCGCTGTTCTGTGCTATAGCCTTAAAA	5100
Db	16731	AAAAAGCTACTCAGGAAACCACTGCGCTGCTGTTCTGTGCTATAGCCTTAAAA	16672
Oy	5101	AGAACTTGCAACTTTCCTCCAAAAAGACAGACTTACGCCATACGAAATGCTGTATGGAGG	5160
Db	16671	AGAAATACAAATCTCCCCCAAAAAAGACAGACTTACGCCATACGAAATGCTGTATGGAGG	16512
Oy	5161	CCCTTCATTAACCATGACCTTGTGCTTGAACCA-AGAAGCCAACTTATGTCAGACATC	5219
Db	16611	CCCTTCCTAACCAATGACCTTGTGCTTGAACCGAGAGCGCCAAATTAGTGTGACATAC	16552
Oy	5220	ACCTCCTTAGCAATATATCAACAGTCTTAAAAATTAACAAGAACTTACCTCAGAA	5279
Db	16551	ACCTCCTTAGCAATATATCAACAGTCTTAAAAATTAACAAGAACTTGTACCCAGAG	16492
Oy	5280	GAGGAAAAAGAACTATTCCACCCMWGTGACATGATATTAGTCAAGTCCCTCTCTAAT	5339
Db	16491	GAGGAAAAAGAA--ATTCCACCCMGATACATGATATTAGTCAAGTCCCTCTTAAG	16435
Oy	5340	TCCCCATCCTTAGATATCACTCTGGGAAAGACCTTACCCAGTCACTTTATTTACCCCACT	5399
Db	16434	TCCCCACCCCTAGATATCACTATCTGGGAAAGACCCCTACCAATCTTTATCTTACCCGAGT	16375
Oy	5400	GGCGTTAAAGGGCTGGAGTGGAGATCTTGGATATCATCACCTTGAGTCAAAATCCTGGATA	5459
Db	16374	GGAGTTAAAGTGTCTGGAGTGGAGTCTTTAGACACA-----TCAAACTCTGGATA	16326
Oy	5460	CTGCCAAAGAACTGTAAAAATCCAGAGACAAACGCTATGCTATTCCTGTGAACCTCTAGAG	5519
Db	16325	CCGCCAAAGAAACCCAAAAATTCAGAGACAAACGCTATGCTATTCCTGTGAACCTCTAGAG	16266
Oy	5520	GATTTGCGCTGCTCTTCAACAAACAACAGAGAAAGTAACTAAATATATAATCCCC	5579
Db	16265	GATCTGACACTGCTCTTCAACCGACAACCTGGAGAAAGTAACTGAATGTGAGACTCC	16206
Oy	5580	CATGAGCCTCCCTTATCATTTTTTCTCTKTAAGTSTTTVA-CCCTSTTTACACTCA	5638
Db	16205	--ATGAGCCCCCGCTCATATTTTTCTCTTATACGTGTGCTTACCCCCCTTACATATCA	16148
Oy	5639	CTGCAACCCCTTCACATGCGCTGTATGACACAGTACCTCCCTVACMAAGATTCTTATGGA	5698
Db	16147	CTCAACCCCTTCATGCGCTATATCTACAGTACTCCCTTACAAAGACTCTTATGGA	16088
Oy	5699	GAATGACAGCTCCCGAAATATATGATGCCCATCTGTATAGAGATCTTSTTAAAGGAACCC	5758
Db	16087	GAATGCGGCTTCCCAAGAAATATATGATGCCCATCTGTATAGAGATCTTSTTAAAGGAACCC	16028
Oy	5759	CAACTTATCTGCCACACCCATATGCCCCGCAACTGCG	5796
Db	16027	CAACTTATCAATCAACCCATATGCCCCCTGCACCTTC	15990

RESULT 4
US-09-949-016-12249/C
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C0001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 12249
? LENGTH: 77997
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: (1) _ (77997)
? LOCATION: (1) _ (77997)
? OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

Query Match: 39.5%; Score 2994.2; DB 4; Length 77997;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;
```

QY	1812	ACCTTGAGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATGAAATGCTATAGGCGTTCCT	1871
Db	19935	ACCTTGAGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATGAAATGCTATAGGCGTTCCT	19876
QY	1872	TCCAGTGCGGTCTACAGAGCACCTTTAAAAAGATGTCCTCAAGTGAAGTAAACCGCCCC	1931
Db	19875	TCCAGTGTGGTCTACAGAGCACCTTTAAAAAGATGTCGATGAATGAATTAAGCTGCCCC	19816
QY	1932	TTGCGTCATGCGCCCTTAATTTCAAGGGAATCACTGGAAGGCGCACCTGCCACAGGGGACAA	1991
Db	19815	-TATCTCAATGCCCTCATGCTCAAGGGAATCACTGGAAGGCGCACCTGCCACAGGGGATGA	19755
QY	1992	GGTCTTTTGAATCAGAAAGCACTAACCAATGATTCAGACGACAGACCTGAGGGTGCCTGG	2051
Db	19756	GGTCTCTTTCAGTACAGAAAGCCACTTAACCAATTAAGCCACGACGAGACCTGAGGGTGCCTGG	19697
QY	2052	GGGAAGGCGCATCCCATGCAATCAACCTCAGAGAGCCCTGGGTATGCTTGAACATTGAGG	2111
Db	19696	GGCAAGGCGCACCCCAT-----GCCCGAGTAATGCTTGACCAATTGGGG	19654
QY	2112	GCCAGGAGGT---TGTCTCTTGAACACTGTGCGGTCTTCTTAATCTTACTCTTCTGTG	2168
Db	19653	GCCAGAGGTAACTGTCTCTCTGGAACACTGGCATGCGCTTCTCAAGCTTAAGCTCTGCTG	19594
QY	2169	CCGGAACACTGTCTCTCAAGTCTGTCACTAATTTCTGAGGGGGTCCMTAAGACGGGAGTCA	2228
Db	19593	CTGAGCAACCGTCTCCCAATCTGTCACTAACC--GAAGGGTCTCAAGATAGGCAGTCA	19537
QY	2229	CTAAGTACTTCTTTCGCCAGCCACTAAGTATGAACTGGGAGGCTTATCTTTCATATG	2288
Db	19536	CTAAGTATC-TTCTCCAGCACTAAGTGTG-ACGTGGGGAATTATCTTTCATCTTGC	19479
QY	2289	TTTTCTAATATAGCTTGAAAGCCCACTACCTTTGTAAGGAGAGACATTTAGCAAAAGC	2348
Db	19478	CTTTCTAATATATCTCGAAAGCCCAACACCTTTGTAAGGGAATACATTTAGCAAAAGC	19419
QY	2349	AGGGGCGCATATACACTCGTAACATAGAGAAAGAAACACCGTGTGTGTGTCCTGCTTG	2408
Db	19418	AGGGGCGCATATATACACTCTGAACATGGGAAAGAAACACCGTTTGCT--GTCCCTTACTTG	19361
QY	2409	AGAAAGGAATTAATCTCTGAAGTCTGGGCAACGAAGGACAAATATGGACGAGCCAAAGAT	2468
Db	19360	AGGAAGGAATTAATCTCTGAAGTCTGGGCAACGAAGGACAAATATGGATAG--GGAAGAT	19302
QY	2469	GCCGCTCTGCTCAACTTAACCTAAAGATTCACCTTCCCTCAAGAAAGCAAGTAC	2528
Db	19301	GCCGCTCCGCTCAAGTAACTAAAGATTCGCTCTTTCCTCCCAAGAAAGCAAGTAC	19242

Db	17095	AAACAAGAAACGGGCGATTATCCACAGAAAGACCTGGCAACTGATTTTACCCACATGCCCA	17036
Oy	4745	AACCTCAGGGATTTTCAGATATCTAAGTCTGGGATATATCTTACACGGGTGGCCAPAGG	4804
Db	17035	AATCTCAGGATTTTCAGATATCTAATCTCGATAGATATCTTTCACCTGGTGGGAGAGG	16976
Oy	4805	CCTTTCCCTGTAGACAGAAAAAGGCCCAAGAGTAAATMAAGCACTAAGTTCAATGAATAA	4864
Db	16975	CCTTTCTCTGTAGACAGAAAAAGGCCCAAGAGTAAATMAAGCACTAATTCAGAAATAA	16916
Oy	4865	TTCCCAATTTGGGACTTCCCGGAGCTTACAGATGCAATPACCCGCTTCCAGAGCA	4924
Db	16915	TTCCCAATTTGGGACTTCCCGGAGCTTACAGATGCAATPAGCCCCGCTTCCAAAGCTG	16856
Oy	4925	CAGTAAACCCAGGAGATATCCAGAGCGG----TTAGGTATACGATATCACTTACACTGCGCC	4980
Db	16855	CAGTAAACCCAGGAGATATCCAGAGCGTATCTTGGCATACAAATATCACTTACACTGCGCC	16796
Oy	4981	TGAAGGCCACAGTCTCTCAGGGAAGTGCAGAAAAATGAATGAAYACTCAAAGACATCTTA	5040
Db	16795	TGAAGGCCACATCTCTCAGGGAAGTGCAGAAAAATGAAGAAACACTCAAACATCTTA	16736
Oy	5041	AAAAAGCAAAACCCAGGAAACCCCACTCAATGCGCTGTCTGTGGCTATAGCCTTAA	5100
Db	16735	AAAAAGCTAACTCAGGAAACCCCACTCGCATGCGCTGTGTGGCCATAGCCTTATTA	16676
Oy	5101	AGAAATCTGCAACTTTCCCAAAAAAGCAAGGACTTACGCCATACGAAATGCTGTAGGAAG	5160
Db	16675	AGAAATCAAAACTCCCCCAAAAAAGCAAGGACTTACGCCATACGAAATGCTGTAGGACAG	16616
Oy	5161	CCCTTCATACCAATGACCTTGTGCTTGACCCA--AGAACAACCACTTAAGTTCAGACATC	5219
Db	16615	CCCTTCCTAACCAATGACCTTGTGCTTGACCGAGACAGGCAATTTAGTTCAGACATC	16586
Oy	5220	AACCTCTTAGCCAAATATCAACAAGTTCTTAAAAATTACAAGAAACCTTACCTGAGAA	5279
Db	16555	AACCTCTTAGCCAAATATCAACAAGTTCTTAAAAATTACAAGGAACTGTATACCGAGAG	16496
Oy	5280	GAGGGAAGAAAGACTATTCACCCMWGTGACATGATATTAGTCAAGTCCCTTCCTGAAT	5339
Db	16495	GAGGGAAGAAAG--ATTTCACCCCTGTGTACATGATATTAGTCAAGTCCCTTCCTGAAG	16433
Oy	5340	TCCCATCCCTAGATACATCTCTGGGAAAGAACCTTACCAGTCACTTTATATTAACCCCACT	5399
Db	16438	TCCCATCCCTAGATACATCTGGGAAAGAACCTTACCAGTCACTTTATCTACCTGAGT	16378
Oy	5400	GCGGTTAAAGTGGCTGAGATGAGTCTTGGATACATCACCTGAGTCAAAATCTTGAT	5459
Db	16378	GCAGTTAAAGTGTCTGAGATGAGTCTTGAACACA-----TCAAAACCTGGAT	16333
Oy	5460	CTGGCAAAAGAACTGAAATTCAGAGAACAAACGATGCTATTCCTGTGAACCTTGAAG	5519
Db	16339	CCGCAAAAGAACTGAAATTCAGAGAACAAACGATGCTATTCCTGTGAACCTTGAAG	16277
Oy	5520	GATTTCGCGCTGCTCTTCAAAACAACAACAGAGGAAAGTAACTAAATCATTAATCCCC	5579
Db	16269	GATCTGCACCTGCTCTTCAAGGACAACTGGAGAAAGTAACTGAAATGATGAGTCC	16211
Oy	5580	CATGSCCTCCCTTATCATATTTTCTCTKATGTTSTTTA--CCCTSTTCACTCTCA	5638
Db	16209	--ATGGCCCCCCCCCTCATATTTTCTCTTACTGTTGTCTTACCCCCCTTTCACATCA	16155
Oy	5639	CTGCACCCCCCTCATAGCCGCTGTATGACACAGTACGTCCCTTCACAGAGTTCTTAGA	5698
Db	16151	CTGCACCCCCCTCATAGCCGCTATATACAGTACGTCCCTTACCAAGAGCTTITAGA	16093
Oy	5699	GAAATCAGAGTCCCGGAAATATTTAGTCCCATCTATAGAGTCTTTSTTAAGGAACCC	5758
Db	16091	GAAATGAGTCTCCCAAGAAATATTTAGTCCCATCTATAGAGATTTTCTTAAGGAACCC	16033
Oy	5759	CCACCTTCACTGCCACACCCATATGCCCGCACTGC	5796

```

DB              16031 CCACTTCACATCCACACCACTATGCCCTGCACTTC 15994

RESULT 5
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312821
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genentech Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B,AJ172A
; CURRENT APPLICATION NUMBER: US//09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3

```

Query Match	38.1%	Score 2888.6	DB 3	Length 2946
Best Local Similarity	98.2%	Pred. No. 0		
Matches 2872	Conservative 49	Indels 3	Gaps 1	
QY 4658	CTGCCTTATGGCCAAAGCTCTTCAAGGAAACAAARAAACAGGCCATTAACCTGARAARAC	4717		
DB 6	CTGCTTATGCCCAAGCTCTTCCAGAGAAACAAAGAACGGCCATTACCTCGAGAAAGAC	65		
QY 4718	TGGCAACTGATTTTAAACCAAGGCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG	4777		
DB 66	TGGCAACTGATTTTAAACCAAGGCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGG	125		
QY 4778	TAAATACTTTCAAGGGTTGGGCAAGGCTTCCCTGAGAGACAGAAAGGCCCAAGAG	4833		
DB 126	TAGAACTTTCAAGGGTTGGGCAAGGCTTCCCTGAGAGACAGAAAGGCCCAAGAG	185		
QY 4838	TAAATAAGGCACTAGTTGATGAAATTAATTCAGATTGGAATTCCCGAGGCTTACAGA	4897		
DB 186	TAAATAAGGCACTAGTTGATGAAATTAATTCAGATTGGAATTCCCGAGGCTTACAGA	245		
QY 4898	GTGACNATAGCCCTGCTTTCCAGGCCACAGTAAACCCAGGAGTATCCAGGCCTTAGTTA	4955		
DB 246	GTGACNATAGCCCTGCTTTCCAGGCCACAGTAAACCCAGGAGTATCCAGGCCTTAGTTA	305		
QY 4958	TACGATATCACTTAACTGCGCGCTGGAAGGCCACAGTCTCCAGGGAAAGGTGAGAAATGA	5011		
DB 306	TACGATATCACTTAACTGCGCGCTGGAAGGCCACAGTCTCCAGGGAAAGGTGAGAAATGA	365		
QY 5018	ATGAATAVCTCAAAAGGACATCTAAAAAAAGCAAAACCCAGAAACCCACCTCACATGGCCTG	5077		
DB 366	ATGAATAVCTCAAAAGGACATCTAAAAAAAGCAAAACCCAGAAACCCACCTCACATGGCCTG	425		
QY 5078	YTCGTGTGCTATAGCCTTAAAAAGATCTGCACTTTTCCCAAAAAGCAGACCTTAGCC	5133		
DB 426	CTCTGTGTGCTATAGCCTTAAAAAGATCTGCACTTTTCCCAAAAAGCAGACCTTAGCC	485		
QY 5138	CATAGGAATGCTGTAATGGAAGGCCCTCATTAACAATGACCTTGTGCTGACCCCAAGAC	5197		
DB 486	CATAGGAATGCTGTAATGGAAGGCCCTCATTAACAATGACCTTGTGCTGACCCCAAGAC	545		
QY 5198	AGCCAACTTAGTTGGACATCACTCTTACCCAAATATCAACAAGTTCTTAAAAATT	525		
DB 546	AGCCAACTTAGTTGGACATCACTCTTACCCAAATATCAACAAGTTCTTAAAAATT	605		

QY	5258	ACAAAGAACCTTATCCGGAGAAAGGAGAAAGAAACATATCCACCCMGMGAGATAGTAT	5317
Db	606	ACAAAGAACCTTATCCGGAGAAAGGAGAAAGAAACATATCCACCCCTTGGAGATGGTAT	665
QY	5318	AGTCAAGTCCCTTCCTCAATTCCTCCATCCCTAGATCATCTCGTGAAGACCTTACC	5377
Db	666	AGTCAAGTCCCTTCCTCAATTCCTCCATCCCTAGATCATCTCGTGAAGACCTTACC	725
QY	5378	AGTCATTTATTTATCCCAACTGCGGTAAAGTGGCTGAGTGAAGTCTTGGATACATCA	5437
Db	726	AGTCATTTATTTATCCCAACTGCGGTAAAGTGGCTGAGTGAAGTCTTGGATACATCA	785
QY	5438	CACCTGAAGCAAAATCCTGGATATCTGCCAAAGAACCTGAAATCCAGAGAACAAAGCTAG	5497
Db	766	CACCTGAAGCAAAATCCTGGATATCTGCCAAAGAACCTGAAATCCAGAGAACAAAGCTAG	845
QY	5498	CAATTCCCTGTAACCTCTAGAGAAATTTGGCGTCTTCAAAACAACAACGAGAGAAA	5557
Db	846	CAATTCCCTGTAACCTCTAGAGAAATTTGGCGTCTTCAAAACAACAACGAGAGAAA	905
QY	5558	GTAACATAAATCATTAATCCCCCATGAGCCCTCCCTATCATATTTTCTCTTAACTGTTT	5617
Db	906	GTAACATAAATCATTAATCCCCCATGAGCCCTCCCTATCATATTTTCTCTTAACTGTTT	964
QY	5618	TTTTAACCTCTTTCACCTCTCACTGCAACCCCTTCATGCGCGCTGTATGACACAGTACCTCC	5677
Db	965	TTTTAACCTCTTTCACCTCTCACTGCAACCCCTTCATGCGCGCTGTATGACACAGTACCTCC	1024
QY	5678	CTAACCMAGAGTTTCTATGGAATGACAGCGTCCCGGAAATTTGATGCCCATGCTATA	5737
Db	1025	CTAACCMAGAGTTTCTATGGAATGACAGCGTCCCGGAAATTTGATGCCCATGCTATA	1084
QY	5738	GGAGTCTTTTSTAAGGAAACCCCACTTCACTGCCACACCCATATGCCCCGCACTGCT	5797
Db	1085	GGAGTCTTTTSTAAGGAAACCCCACTTCACTGCCACACCCATATGCCCCGCACTGCT	1144
QY	5798	ATCACTCTGCGCACTCTTTCGATGATGACATGCAATATCTATTTGACAGAGAAAATGATTA	5857
Db	1145	ATCACTCTGCGCACTCTTTCGATGATGACATGCAATATCTATTTGACAGAGAAAATGATTA	1204
QY	5858	ATCCAGTGTCTCTGGAGACTTGAAGTCACTGCTGTGGACTTACTTCCACCCCAACTG	5917
Db	1205	ATCCAGTGTCTCTGGAGACTTGAAGTCACTGCTGTGGACTTACTTCCACCCCAACTG	1264
QY	5918	GTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAAGAGTAA	5977
Db	1265	GTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAAGAGTAA	1324
QY	5978	TCTCCCACTCACGSGGGGTACATGCGACCTCTAGGCCCTTCAAAAGACTAGATCTCTCA	6037
Db	1325	TCTCCCACTCACGSGGGGTACATGCGACCTCTAGGCCCTTCAAAAGACTAGATCTCTCA	1384
QY	6038	AACTACATGAAAACCTTCCTGTAACCATATCTCGCGCTGTAAAGCTATTTTAATACACCTCA	6097
Db	1385	AACTACATGAAAACCTTCCTGTAACCATATCTCGCGCTGTAAAGCTATTTTAATACACCTCA	1444
QY	6098	CTGGAGCTCCATGAGGTCTCGGCGCAAAAACCTACTACTGTTGATATGCTCCCTCGTA	6157
Db	1445	CTGGAGCTCCATGAGGTCTCGGCGCAAAAACCTACTACTGTTGATATGCTCCCTCGTA	1504
QY	6158	ACTTGARGCCATATGTTTCAATCCTGTATCTGAACTTGAACAACTTCAGACAGAAA	6217
Db	1505	ACTTGARGCCATATGTTTCAATCCTGTATCTGAACTTGAACAACTTCAGACAGAAA	1564
QY	6218	TAAACACCACTTCCTGTTTATGAGACCTCTTGTGTTTCCAAATSTGAAATTAACCTTACT	6277
Db	1565	TAAACACCACTTCCTGTTTATGAGACCTCTTGTGTTTCCAAATSTGAAATTAACCTTACT	1624
QY	6278	CAAACTCAACCTGTAAATTTTGAATACATACATACAAACCAACTCCCAATGATCATCA	6337
Db	1625	CAAACTCAACCTGTAAATTTTGAATACATACATACAAACCAACTCCCAATGATCATCA	1684
QY	6338	GGTGGGTAACTCTCCCAACAAATAGCTGCTCACTCCTCAAGAAATATTTTGTCTGTG	6397
Db	1685	GGTGGGTAACTCTCTCCCAACAAATAGCTGCTCACTCCTCAAGAAATATTTTGTCTGTG	1744
QY	6398	GTAACCTCAGCCTATTCGTTGTTTGAATGAGCTCTTCAGATCTATGCTCTCTCATCT	6457
Db	1745	GTAACCTCAGCCTATTCGTTGTTTGAATGAGCTCTTCAGATCTATGCTCTCTCATCT	1804
QY	6458	TAGTGCCCCCYATGRCATTTACATGAAACAATTTTATACATGTTATGTCAATCTAAGC	6517
Db	1805	TAGTGCCCCCYATGRCATTTACATGAAACAATTTTATCAATTTATGTCAATCTAAGC	1864
QY	6518	CCCGCAACAAAGAGTACCACTCTTCCTTTTGTATAGAGAGAGAGTCTAGTGAC	6577
Db	1865	CCCGCAACAAAGAGTACCACTCTTCCTTTTGTATAGAGAGAGAGTCTAGTGAC	1924
QY	6578	TAGTACTGSCATTTGGGGATACAAACCTCTCACTCAGTTCTCTCAAACTATCTCAAG	6637
Db	1925	TAGTACTGSCATTTGGGGATACAAACCTCTCACTCAGTTCTCTCAAACTATCTCAAG	1984
QY	6638	AACTAAATGGGACATGAAACGGGTCGCGACTCTCCTGTGACCTTGCAGAAATCAACTTA	6697
Db	1985	AACTAAATGGGACATGAAACGGGTCGCGACTCTCCTGTGACCTTGCAGAAATCAACTTA	2044
QY	6698	ACTTCCCTAGACAGATGATCTCTTCAAAATGAAAGACTTTAAGCTGTCAACCGCTGABA	6757
Db	2045	ACTTCCCTAGACAGATGATCTCTTCAAAATGAAAGACTTTAAGCTGTCAACCGCTGABA	2104
QY	6758	GAGGGGAACTGTATATTTTATGGGAAAGAAATGCTGTATATTTATCAATCCGAA	6817
Db	2105	GAGGGGAACTGTATATTTTATGGGAAAGAAATGCTGTATATTTATCAATCCGAA	2164
QY	6818	TCGTCACTGAGAAAGTTTAAAGAAATTCAGATGAAATACAACTGTAACAGAGAGCTTC	6877
Db	2165	TCGTCACTGAGAAAGTTTAAAGAAATTCAGATGAAATACAACTGTAACAGAGAGCTTC	2224
QY	6878	GAAACACTGGAACCTTGGGGCTCTCTCAAGCCTTCAAGATGATGCTGCTGCTCTTCTAG	6937
Db	2225	GAAACACTGGAACCTTGGGGCTCTCTCAAGCCTTCAAGATGATGCTGCTGCTCTTCTAG	2284
QY	6938	GACCTCTAGAGCTATATATTTGCTACTCCTCTTGGACCTGTATCTTTTACCTCTTG	6997
Db	2285	GACCTCTAGAGCTATATATTTGCTACTCCTCTTGGACCTGTATCTTTTACCTCTTG	2344
QY	6998	TTAACTTGTCTCTTCCAGAAATGGAAGCTGTAAACTTACAAATGAGCCCAAGATCAGT	7057
Db	2345	TTAACTTGTCTCTTCCAGAAATGGAAGCTGTAAACTTACAAATGAGCCCAAGATCAGT	2404
QY	7058	CCAAAGCTAAGATCTACCGGACACCCCTGGAACCGGCTGTAGCCCAAGATCTGATGTTA	7117
Db	2405	CCAAAGCTAAGATCTACCGGACACCCCTGGAACCGGCTGTAGCCCAAGATCTGATGTTA	2464
QY	7118	ATGACATCAAAAGGACCCCTCCGAGAAATCTCAGCTGACAAACCTTACTAGGCCCCA	7177
Db	2465	ATGACATCAAAAGGACCCCTCCGAGAAATCTCAGCTGACAAACCTTACTAGGCCCCA	2524
QY	71		

Db 2765 AGAATATGCCATCATCTATTGCTGAGACACAGAGGAGCAATGATCGGATATA 2824

Qy 7478 AACCCAGTATTGAGCGGACGAGCAACCCCTTTGGGTCCCTCCCTTTGATGGGA 7537

Db 2825 AACCCAGTCTTCAGCGCGGACGAGCAACCCCTTTGGGTCCCTCCCTTTGATGGGA 2884

Qy 7538 GCTCTGTTTCATGCTATTTCATCTTAAATCTTGACCTGCR 7582

Db 2885 GCTCTGTTTCATGCTATTTCATCTTAAATCTTGACCTGCA 2929

RESULT 6

US-09-949-016-13002/c

/ Sequence 13002, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTNER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ FILE REFERENCE: C1001307

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/231,498

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FASTSEQ for Windows Version 4.0

/ SEQ ID NO 13002

/ LENGTH: 168394

/ TYPE: DNA

/ ORGANISM: Human

/ FEATURE:

/ NAME/KEY: misc_feature

/ LOCATION: (1)..(168394)

/ OTHER INFORMATION: n = A,T,C or G

US-09-949-016-13002

Query Match 34.5%; Score 2615.4; DB 4; Length 168394;

Best Local Similarity 85.1%; Pred. No. 0; Mismatches 355; Indels 58; Gaps 33;

Matches 3133; Conservative 134; Mismatches 355; Indels 58; Gaps 33;

Qy 1012 TTCCCAATTATGACCCCTCCCAAGCACTGGAGAGAGAAATTCGCCACGACAGATG 1071

Db 2905 TCCCCCAAGGCAAAAACCCCTTAAATGATTTCTGAGAAATTCGGCTGTGAGATG 28946

Qy 1072 CATGTGCTTTTCTCTCCAGACTTAAAGCAATAAACAGACTTAGTAAATTCGA 1131

Db 28945 TATGTACTTTTCCCTCTGACACTTGAAGCAATTAATAATGACCTAGTAAATTCGA 28886

Qy 1132 GATAACCTGATGCTATATGATGTTTAAACAAGGTTAGGCAATCTTGATCTGACA 1191

Db 28885 GATAACCTGTTGGCTATATGATGCTTTTAAAGGTTTGAACAATCTTGATCTGACA 28826

Qy 1192 TGGAGAGATATATATGCTACTGCTAAATCAGACACTTAAACCCCAATGAGAGAGTCCAC 1251

Db 28825 TGGAGAGATATATATGCTACTGCTAAATCAGACACTTAAATCCCAATGAGAGAGTCCAC 28767

Qy 1252 CATATCTGACCTGAGGTTTGGCGATCTCTGATATCTCACTCAGTCAATGATTAAG 1311

Db 28766 CATATCTGACCTGAGGTTTGGCGATCTCTGATATCTCACTCAGTCAATGATTAAG--ATAGG 28709

Qy 1312 ATGCAATAGAAAGAAATGATTCCTCCCAAGGCAAGGAGTTCCAGTCTGASA 1371

Db 28708 ATGCAATAGAAAGAAATGATTCCTCCCAAGGCAAGGAGTTCCAGTCTGASA 28650

Qy 1372 CCCCTATTGGGAGACAGAAATCACTAAATGAGAGATTTGCTGACAGACATTTGCTAA 1431

Db 28649 CCCCTATTGGG-----ACACAGAAATCAAGAAATGAGATTTGCTGACAGACATTTGCTAA 28595

Qy 1432 CTTGTGTCTAAGAGACTTAAGAAATCTASGAAAGAAATCTAATGATG 1491

Db 28594 CTTGTGTCTA--AAGACTAAGAAATCTAGAAAG--AAGCTTATGATTTTCAATGATG 28537

Qy 1492 TCCACATTAACACAGGAGAAAGGAAATTCCTACTGCTTTCTGAGAGACTTAAGGA 1551

Db 28536 TTCACATTAACACA--AGAAAGAAATTCCTACTGCTTTCTGAGAGACTTAAGGA 28478

Qy 1552 GGCATTTGAGAGAGCGTCTCTCTGTCACCTGACTCTTCTGAAAGCCCACTTAATCTTAA 1611

Db 28477 GGCATTTGAGAGAGCG--AAACTCTGTCTACCTGACTGTTGAAAGCCCACTTAATCTTAA 28419

Qy 1612 GCGTAAGTTTATCACTCACTGACTGACTGACACTTAA--AAAACTTCAAAAGTCTGCGT 1670

Db 28418 GGTTAAGTTTATCACTCACTGACTGACTGACACTTAAAGAAAACTTAAAGTCCACTT 28359

Qy 1671 AGGCCCGAGCAAAACTTAGAAACCTTATTGAACTTGGCACTTCGTTTTTTTAAATAG 1730

Db 28358 AGGCCCGAGCAAAACTTAGAAACCTTATTGAACTTGGCACTTCGTTTTTTTAAATAG 28299

Qy 1731 AGATCAGAGAGACAGCGGAAACAGACAAACGGGATTAATAAAAGCCACGCTTAG 1790

Db 28298 AGATCAGAGAGACAGCGGAAACAGACAAACGGGATTAATAAAAGCCACCTTAG 28240

Qy 1791 TCATGACCCCTCAGCAAG--TGACTTTGAGGCTCTGAAAGGAAAGCTGGCAAT 1849

Db 28239 TCATGACCCCTCAGCAAGCAAGCTTCGAGGCTCTGAAACAGGAAAGCTGGCAAT 28180

Qy 1850 TGAATGCTTAATAGGCTTCTTCACTGCTGCTCAAGAGACACTTTAAAAAGATTTGT 1909

Db 28179 TGAATGCTTAATAGGCTTCTTCACTGCTGCTCAAGAGACACTTTAAAAAGATTTGT 28120

Qy 1910 CCAATGTAAGTAAGCC--GGCCCTTGCTCATGCGCCCTTATTTAAAGGAATCACTGGA 1968

Db 28119 TCAATGTAATGTAAGCC--GGCCCTTGCTCATGCGCCCTTATTTAAAGGAATCACTGGA 28060

Qy 1969 GGCCCACTGCCCCAGGAGCAAAAGTCTTTTGAATCAGAAAGCACTAACAGATGATCA 2028

Db 28059 GGCCCACTGCCCCAGGAGCAAAAGTCTTTTGAATCAGAAAGCACTAACAGATGATCA 28000

Qy 2029 GCAGCAGACTGAGAGTCTGCTGGGCAAGGCGCATCCCATGCTACACCTTCAAGAGCC 2088

Db 27999 GCAGCAGACTGAGAGTCTGCTGGGCAAGGCGCGCCAGCCCATGCTACACCTTCAAGAGCC 27940

Qy 2089 CTGGTATGCTTGAACCTTGAAGGCGCGAAAGGT---TGCTCTGGAACACTGGTGGCT 2145

Db 27939 CTGGTATGCTTGAACCTTGAAGGCGCGAAAGGT---TGCTCTGGAACACTGGTGGCT 27880

Qy 2146 CTTCTAGTCTTACTCTTCTGCTCCGAGCAACTGCTCCAGATCTGCTACTATTTCTAG 2205

Db 27879 CTTCTAGTCTTACTCTTCTGCTCCGAGCAACTGCTCCAGATCTGCTACTATTTCTAG 27821

Qy 2206 GGGGTCCNTAAGACGGGCACTCACTAATCTTTTCCAGCCCTAAGTTATGAATGG 2265

Db 27820 GG-----TCTGAGGACAGCCACTAGATAC--TTCCTCCAGCCACTAAGCTGTG--ACTGG 27770

Qy 2266 GAGCTTATTTCTTTTCAACATGCTTTCTAATATGCTGGAAGCCCACTACCTGTTA 2325

Db 27769 GGAATTTACTCTTTTCAACATGCTTTCTAATATGCTGGAAGCCCACTACCTGTTA 27710

Qy 2326 GGGAGAGACATTTAGCAAAAGCAGGGGCACTTATACCTGAACTATGAGAAAGAAACA 2385

Db 27709 GGGAAAGACATTTCAACAAAGCAGGGGCACTTATACCTGAACTATGAGAAAGAAACA 27650

Qy 2386 CCCGTTTGTGTCCTGCTTGAAGAAAGAAATTAATCTGAAGTCTGGGCAACGAAAG 2445

Db 27649 CCCGTTTGTGTCCTGCTTGAAGAAAGAAATTAATCTGAAGTCTGGGCAACGAAAG 27591

Qy 2446 ACAATATGAGACGAGACCAAGATGCGCTGCTTCAAGTTAACTAAGATTCACATT 2505

Db 27590 ACAATATGAGACGAG--CAATATATGCGCTGCTTCAAGTTAACTAAGATTCACATT 27552

Qy 2506 CTTTCCCTAACAAAGGCAATACCCCTCAGACCCCAAGGCCCAAGAGATTCAAAGGA 2565

Db 27321 CTTTTCCCTACCAAGGAGTACCCCTTAGACCCGAGGCCCAACAGGATCCAAAAA 27472
QY 2566 TTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAACCATGACATACTCCCTCCAGTAATT 2625
Db 27471 TTGTTAAGGACTTAAAGCCCTTACTCTAGTAAACCATGACATAAGCCCTGCAATCCCTC 27412
QY 2626 CCGTAGTGATTTAGAGGAGCAGAAACCAGTGCAGATGAGGAGTTAGTCAAGATCT 2685
Db 27411 C-----AATTTAGAGTAAAGAACCCATAGACGTGGA -GTTTAGTACAAAGAT 27360
QY 2686 CAGGATTTCATGAGAGGCCGTTCCTTTATACCAGCTGTACCTTAGCCTTAGTCTG 2745
Db 27359 CAGGATTTACAGT -GAGCTGTCTTCTCTATACCCAGCTGTACCTTAGCCTTAGCTC 27301
QY 2746 TGMATTTCCCAATATACAGAGGAGAGAGTGGTTTAACTCACTCTGAGCCTTAAAGATGCT 2805
Db 27300 TGCTTTCCCAATATACAGAGGAGAGAGTGGTTTAACTCACTCTGAGCCTTAAAGATGCT 27241
QY 2806 TCTTCTGCATCCCTGTATACCTGTCACTCTCAATTTCTGTGGCTTTGAAGATTTGCA 2865
Db 27240 TTTTCTGCATCCCTGTATACCTGTCAATTTCTGTGGCTTTGAAGATGCTTCA 27181
QY 2866 ACCGATCTCACTCACTGAGACTTTTATCCCAAGGTTCAAGGATAGTCCCATC 2925
Db 27180 ACCCAATGCTCAACTCACTGAGACTTTTATCCCAAGGTTAAGATAGCCCATC 27121
QY 2926 TATTTGGCAGGCAATTAGCCCAAGCTGAGCAATWTTCATCTACCTGAGACTCTGTGCC 2985
Db 27120 TATTTGGCAGGCAATTAGCTCAAGACTGAGCAAGTCTCATCTGAACTGAACTGTGCC 27061
QY 2986 TTCTGATAGTGAATTAATTTTGTGTCGTCCT -CAGAACTTGTGCATCAAGC 3044
Db 27060 TTTGTGATAGTGAATTAATTTGTGTCGTCCTGTTGAAGAACTTGTGCATCAAGC 27001
QY 3045 ACCCAAGCTCTTMAATTTCTGTCGTAAGCTGTGCTACAGGTTTCCAAACATGCT 3104
Db 27000 ACCCAAGCTCTTMAATTTCTGTCGTAAGCTGTGCTACAGGTTTCCAAACATGCT 26941
QY 3105 CACTCTGCTCAGAGGCTTAAATTAATTAAGCAATTAATTAAGCAATTAATTAAGCA 3164
Db 26940 CAGCTCTGCTTACAGAGGCTTAAATTAATTAAGCAATTAATTAAGCAATTAATTAAG 26881
QY 3165 CTCAGTGAAGAAATATCAGGCTTATCTGCTTATCTCATCTCAATCAAAACCTTAAAGCA 3224
Db 26880 CTCAGTGAAGAAATATCAGGCTTATCTGCTTATCTCATCTCAATCAAAACCTTAAAGCA 26821
QY 3225 CTAAAGRRRTTCTTTGGCTTAATVAGTTCCTGCGAAMATGATTTCCAGGTTGCGA 3284
Db 26820 ATAAAG-GTTCCTTGGCATTAACAGGCTTTCGCAAAATATGATTT-CCAGGTACGGCA 26763
QY 3285 AATAGCCAGGCAATTAAATACATTAATTAAGAAATCAAGAAACCAATACCCATTAAT 3344
Db 26762 AATAGCCAGGCAATTAAATACATTAATTAAGAAATCAAGAAACCAATACCCATTAAT 26703
QY 3345 AAGATGAAATCTGAATTAAGTGTCTTTCAGAGCCCTTAAAGAGGCTTAAACCA 3404
Db 26702 AAATGGAATCTTAAGCAAGAAAGAGCTTTCAGAGCCCTTAAAGAGGCTT-AAACCA 26644
QY 3405 GYCCCACTGTTAAGTTCGCAACRGGGCAAGACTTTTSTTAAATVATCAAGAAAAAAC 3464
Db 26643 GACTCAGTGTAAAGCTTAACCAACGGGGCAAGACTTTTCTTCAATATGTCAC--AAAAAC 26587
QY 3465 AGAAATVAGCTCTGGAATCTCTTACAGATCCAGAGGAAATGAGCTTTCGACATCTTGTG 3524
Db 26586 AGAAATVAGCTCTGGAATCTCTTACAGAGGAAATGAGCTTTCGACATCTTGTG 26527
QY 3525 CCTGASTAAGAAATVATGATGATGAGCAAGAGGTTGATCTTATGTTAAGGATGAT 3584
Db 26526 CTTAAGTAAAGAAATTAATGATGAGCAAGAGGTTGATCTTATGTTAAGGATGAT 26467
QY 3585 GGCAGTAGCAGTCTTATCTGAAGCAATTAATTAATTAAGGAAAGATCTTACTGT 3644
Db 26466 GGCAGTAGCAGTCTTATCTGAAGCAATTAATTAATTAAGGAAAGATCTTACTGT 26407

QY 3645 GTGACATCTCATGAKGTGAAYRGCATTAATCACTGTCTTAAAGAGACTTGTGCTGACA 3704
Db 26406 GTGACATCTCATGATGTAATGAACAGCATTAATCACTGTCTTAAAGAGACTTGTGCTGACA 26347
QY 3705 CAACGTTTACTTAAATATCTGAGGCTTATTAATTAAGGAGGAGGCTGACCTGAC 3764
Db 26346 CAATCGTTTGTAAATATCAAGGCTTATTAATTAAGGAGGAGGCTGACCTGACCTGAC 26287
QY 3765 TTGTGCACTCTTAACCCAGCNCATTTCTTCCAGCAATGAAGAAAAATTAATTA 3824
Db 26286 TTGGCAACTCTTAACCCCGCATTTCTTCCAAACATGA--AAAGATTAATTAATTA 26229
QY 3825 CTGTCAACATTAATTTCTGAAATGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3884
Db 26228 CTGTCAACATTAATTTCTGAAATGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 26170
QY 3885 GACTATCTGACCTTCAACTTGTATCTGATGGAAGTTCCTTTGTAGAAAAAGGACTTC 3944
Db 26169 GACTATCTGACCTTCAACTTGTATCTGATGGAAGTTCCTTTGTAGAAAAAGGACTTC 26111
QY 3945 GAAAAATGAGGATGAGTGTGATGATTAAGGAAATATTAAGGATTAATCCCTCACTC 4004
Db 26110 -AAAAAGAGGATGAGTGTGATGATTAAGGAAATATTAAGGATTAATCCCTCACTC 26052
QY 4005 CAGGAATAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4064
Db 26051 CAGGAATAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 25992
QY 4065 AAGAAAAAAGGAAATATATATATACAGACTCTATATATGCTATGCTATGCTATGCT 4124
Db 25991 AAGAAAAAAGGAAATATATATATACAGACTCTATATATGCTATGCTATGCTATGCT 25932
QY 4125 ATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4184
Db 25931 ACACAGCATATATGAAAAAAGGAAATTCCTTAATCTTGAAGGAAACATATCAACATC 25872
QY 4185 AGGAAGCATTAAGAAATTAATTAATGAGGAAATTAAGGAAATTAAGGAAATTAAGGAA 4244
Db 25871 AGGAAGCATTAAGAAATTAATTAATGAGGAAATTAAGGAAATTAAGGAAATTAAGGAA 25812
QY 4245 ACTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4304
Db 25811 ACTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 25752
QY 4305 TTGAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4364
Db 25751 TTGAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 25692
QY 4365 TAGTATAGGATTAATCCCTTCCGGGAAACCAAGCCCAAGTACTGACGAGGAAACAT 4424
Db 25691 TAGTATAGGATTAATCCCTTCCGGGAAACCAAGCCCAAGTACTGACGAGGAAACAT 25634
QY 4425 GGGGAACCTGAC--GAGGCAATTTCTCCCTCGGAGCGGTTAGCCATGAGAAAGGA 4481
Db 25633 GGGGAACCTGAC--GAGGCAATTTCTCCCTCGGAGCGGTTAGCCATGAGAAAGGA 25574
QY 4482 AAATACCTTTGCTGCAACTATCAATGAAATTAATTAATTAATTAATTAATTAATTAAT 4541
Db 25573 AAATACCTTTGCTGCAACTATCAATGAAATTAATTAATTAATTAATTAATTAATTAAT 25514
QY 4542 ACTTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4601
Db 25513 ACTTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 25454
QY 4602 TCAAACTATCAAGCATATCAAGGAGGCTGTGAATGAGGAGGAGGAGGAGGAGGAGG 4661
Db 25453 TCAAACTATCAAGCATATCAAGGAGGCTGTGAATGAGGAGGAGGAGGAGGAGGAGG 25394
QY 4662 CTATGCGCAAGCTCTTCA 4681
Db 25393 ACTGACAGGCAATTAATTTCA 25374

RESULT 7
US-09-949-016-15858/c
Sequence 15858, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
CURRENT FILING DATE: 2000-04-14
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: PASTESEQ for Windows Version 4.0
SEQ ID NO 15858
LENGTH: 145320
TYPE: DNA
ORGANISM: Human
US-09-949-016-15858

Query Match 29.3%; Score 2220.4; DB 4; Length 145320;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 2728; Conservative 92; Mismatches 216; Indels 115; Gaps 32;

QY 606 GATGGGAACGTTCCCGCAGACAAACCGCCCTAGACGTATTCGGAAATTGGGA 665
DB 97137 GATGGGAACGTTCCCGCAGACAAACCGCCCTAGACGTATTCGGAAATTGGGA 97078
QY 666 MCAATTTGACCTCTGACACCTAAGAAAGAAAGCACTTATCTCTGACAGTCCGCTG 725
DB 97077 CCAATTTGACCTCTGACAGTCTAAGAAAGAAAGCACTTATCTCTGACAGTCCGCTG 97020
QY 726 GCA-----CTCTGAGGAGATTAATTATTA 753
DB 97019 GCAACGATATCTCTTAAGGGGAGAAACCTGCATCTGAGGAGATTAATTATTA 96960
QY 754 CACCATCTTAACGCTAGACCTTTTGTGAA--AAGGCAATGAGTGAATGCCCTA 810
DB 96959 CACCATCTTAACGCTAGACCTTTTGTGAAAGAAAGCAATGCTGAAATGCTCA 96900
QY 811 AGTCAAACTTCTTTCACTTAAGAGCAACTCACATTAATTAAGTGTGATTTATG 870
DB 96899 GGTCAAACTTCTTTCACTTAAGAGCAACTCACATTAATTAAGTGTGATTTATG 96840
QY 871 CCTACAGAAAGCTTCAAGAGTCTACCTCCCTATCCAGACATCCCC--GACTCTTCCC 928
DB 96839 CCTACAGAAAGCTTCAAGAGTCTACCTCCCTATCCAGACATCCCCGAGCTTCCC 96780
QY 929 AMTTATTAAGACCCCTTCAACCAATGCTCCAAAGAGATTAAGCAAAAGGTTAA 988
DB 96779 AAATATTAAGACCCCTTCAACCAACGCTCCAAAGAGATTAAGCAAAAGGTTAA 96720
QY 989 CAGTAAACCAAGAGTCCCAATATTCCTCCCAATTATGACCCCTCCCAAGCAGTGGAGAA 1048
DB 96719 CAGTAAACCAAGAGTCCCAATATTCCTCCCAATTATGACCCCTCCCAAGCAGTGGAG 96664
QY 1049 GAGAAATCGCCAGCAGAGTGAATGTCCTTTTCTCCAGACTTAAAGCAATTA 1108
DB 96663 GAGAAATCGCCAGCAGAGTGAATGTCCTTTTCTCCAGACTTAAAGCAATTA 96608
QY 1109 AAACGACTTAAATTTCTGAGATTAATCTGATGCTATATGTTGTTTAAAGGTT 1168
DB 96607 AAATGACTTAAATTTCTGAGATTAATCTGATGCTATATGTTGTTTAAAGGTT 96548
QY 1169 TAGGCAATCTTTATGATGATGAGATTAATGATCTGCTAAATCAGACACTA 1228
DB 96547 TAGGCAATCTTTATGATGATGAGATTAATGATCTGCTAAATCAGACACTA 96489

QY 1229 ACCCAATATGAGAAATGCGCACATTAATGAGAGCTTGGGATCTGCTGAT 1288
DB 96488 ACCCAATATGAGAAATGCGCGTAAATGAGAGCTTGGGATCTGCTGAT 96429
QY 1289 CTGACGAGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1348
DB 96428 CTGACGAGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96372
QY 1349 CAGCAGCACTTCCAGCTTASACCTTATGGGAGACAGAAATCAGTAATGAGGA 1408
DB 96371 CAGCAGCACTTCCAGCTTASACCTTATGGGAGACAGAAATCAGTAATGAGGA 96316
QY 1409 TTGCTGCTGAGACATTTTCTTAACTTGTGCTTAAAGCTAAGAAATAGGAAGA 1468
DB 96315 TTGCTGCTGAGACATTTTCTTAACTTGTGCTTAAAGCTAAGAAATAGGAAGA-A 96257
QY 1469 ARTCTAAGATTAATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1528
DB 96256 AGCTGAGATTAATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96198
QY 1529 GCTTTTGTGAGAGACTTAAGAGGAGCTTGAAGAGGCTGCT--CTGCACTGAT 1586
DB 96197 GCTTTTGTGAGAGACTTAAGAGGAGCTTGAAGAGGCTGCT--CTGCACTGAT 96138
QY 1587 CTCTGAGAGGCACTAATCTTAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 1646
DB 96137 CTCTGAGAGGCACTAATCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 96078
QY 1647 -GAAAAAATCTCAAAAGTCTGCGCTAGAGCCCGAGCAAACTTGAACCTTATGA 1705
DB 96077 GAAAAAATCTCAAAAGTCTGCGCTAGAGCCCGAGCAAACTTGAACCTTATGA 96018
QY 1706 TGCGCACTGCGCTTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765
DB 96017 TGCGCACTGCGCTTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 95958
QY 1766 ATTAAAAAAGGCAACCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1825
DB 95957 A-TAAAAAAGGCAACCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 95899
QY 1826 GAAAAAGGAAAGCTGCGCAATTTGAATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 1885
DB 95898 GAAAAAGGAAAGCTGCGCAATTTGAATGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 95839
QY 1886 CAAGGACCTTAAAAAAGATTTGCTCAAGTAAAGTAAAGGCTTAAAGGCTTAAAGGCTTAA 1945
DB 95838 CAAGGACCTTAAAAAAGATTTGCTCAAGTAAAGTAAAGGCTTAAAGGCTTAAAGGCTTAA 95779
QY 1946 TTATTTTCAAGGAAATCACTGGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 2005
DB 95778 TTATTTTCAAGGAAATCACTGGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 95719
QY 2006 GAGGCACTTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2065
DB 95718 GAGGCACTTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 95659
QY 2066 CATGCACTTACCTTCAAGAGCTTGGGATGCTTGAACCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 2122
DB 95658 CATGCACTTACCTTCAAGAGCTTGGGATGCTTGAACCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 95599
QY 2123 TGTCTCTGAGACATGAGGCTTCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 2182
DB 95598 TGTCTCTGAGACATGAGGCTTCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 95539
QY 2183 TCCAGATCTGCTCACTTATCTGAGGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 2242
DB 95538 TCCAGATCTGCTCACTTATCTGAGGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 95483
QY 2243 CCAGCCACTAAGTATGAATGAGGAGCTTAAATCTTTTCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 2302
DB 95482 CCAGCCACTAAGTATGAG-AGTGGGAACTTCACTTTTCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAA 95424

QY 2203 TTGAAAGCCCACTACCTTGTAGGAGACATTTCTAGCAAAAAGAGGGCCATTATAC 2362
DB 95423 CTGAAGGCCAATCCCTTGTAGGAGAGACATTTCTAGCAAAAAGAGGGCCATTATAC 95364
QY 2263 ACCCTGAATATGAGAGAGAAACCCCTTTGTGTCCTTGTGAGAGAAAGAAATAT 2422
DB 95363 ACCCTGAATATGAGAGAGAAACCCCTTTGTGTCCTTGTGAGAGAAAGAAATAT 95309
QY 2423 CCTGAAGCTGGGCAACGAAAGGACATATGAGAGAGGACCAAGATGCGCTGTTCA 2482
DB 95308 CTGAAGCTGGGCAACGAAAGGACATATGAGAGAG--CAAGAAATGCGCTGTTCA 95250
QY 2483 AGTTAACTAAAGATTCACCTTCTTCCCTACCAAGAGAGATCCCTCAGACCCA 2542
DB 95249 AGTTAACTAAAGATTCAGCTCTTCTTCCCAACCAAGAGAGATCCCTCAGACCCA 95190
QY 2543 GGCCCAACAGAGATTCACCAAGATTTGTTAAGACTTAAAGCCAGGCTTATGAAAAC 2602
DB 95189 GGCTCAACAGAGATTCACCAAGATTTGTTAAGACTTAAAGCCAGGCTTATGAAAAC 95133
QY 2603 ATGATTAATCTCCGCAATATTCGTAAGTGAATGAGAGGACAGAAACCAAGTGGAC 2662
DB 95132 ATGATTAATCTCCGCAATATTCGTAAGTGAATGAGAGGACAGAAACCAAGTGGAC 95080
QY 2663 AGTGAAGGTTAGTGAAGATCTCAGATTAATCAATGAGAGGCGCTTGTATATACC 2722
DB 95079 AGTGAAGGTTAGTGAAGATCTCAGATTAATCAAT--GAGTCACTGCTCCTATACCT 95022
QY 2723 AGCTGACTAAGCTTATCTGATGATTTCCCAATATCAGAGAGAGAGAGAGTGTAC 2782
DB 95021 AGCTGACTAAGCTTATCTGATGATTTCCCAATATCAGAGAGAGAGAGAGTGTAC 94962
QY 2783 AATCTGAGAGCTTGAAGATGCTTCTGATCCTGATCCTGATCCTGATCCTGATCCT 2842
DB 94961 AATCTGAGAGCTTGAAGATGCTTCTGATCCTGATCCTGATCCTGATCCTGATCCT 94902
QY 2843 GTTTCCTTTGAAGATCTTCAAAACCCAGATCTCACTCACTGATCTTATATACCA 2902
DB 94901 GTTTCCTTTGAAGATCTTCAAAACCCAGATCTCACTCACTGATCTTATATACCA 94842
QY 2903 AGGCTGAGAGATGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2962
DB 94841 AGGCTGAGAGATGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 94782
QY 2963 ATGATTAATCTCCGCAATATTCGTAAGTGAATGAGAGGACAGAAACCAAGTGGAC 3020
DB 94781 ATGATTAATCTCCGCAATATTCGTAAGTGAATGAGAGGACAGAAACCAAGTGGAC 94722
QY 3021 TCAGAAACCTTGTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3080
DB 94721 TCAGAAACCTTGTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94662
QY 3081 TACAGGTTTCCAAACGAAAGATCTGATCTGATCAAGAGAGAGAGAGAGAGAGAG 3140
DB 94661 TACAGGTTTCCAAACGAAAGATCTGATCTGATCAAGAGAGAGAGAGAGAGAGAG 94602
QY 3141 AATTAATCAAG 3200
DB 94601 AATTAATCAAG 94556
QY 3201 CCTCATCAAAACCCCTAAAG 3260
DB 94555 CCTCATCAAAACCCCTAAAG 94496
QY 3261 AATATGATTTCCCAAG 3320
DB 94495 AATATGATTTCCCAAG 94437
QY 3321 TCAGAAAGGCAATCCCATTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 3380
DB 94436 TCAGAAAGGCAATCCCATTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 94377
QY 3381 CCTTAAGAT 3440

DB 94376 CTAAAG 94319
QY 3441 TATTAATATATCAAG 3500
DB 94318 TATTAATATATCAAG 94264
QY 3501 GATGAG 3560
DB 94263 GATGAG 94204
QY 3561 GATGAG 3620
DB 94203 GATGAG 94144
QY 3621 AATTAAG 3680
DB 94143 AATTAAG 94088
QY 3681 TAAAG 3711
DB 94087 TAAAG 94057

RESULT 8

US-09-949-016-5640
Sequence 5640, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241, 755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237, 768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231, 498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5640
LENGTH: 2763
TYPE: DNA
ORGANISM: Human
US-09-949-016-5640

Query Match 29.1%; Score 2204; DB 4; Length 2763;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2194; Conservative 35; Mismatches 0; Indels 1; Gaps 1;

QY 5352 GATATCATCTGAG 5411
DB 535 GATATCATCTGAG 594
QY 5412 GCTGAG 5471
DB 595 GCTGAG 654
QY 5472 CTGAAAGATCCAG 5531
DB 655 CTGAAAGATCCAG 714
QY 5532 CTGTAAG 5591
DB 715 CTGTAAG 773
QY 5592 TTAATCATATTTCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 5651
DB 774 TTAATCATATTTCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 833

QY 5652 ATGCCGCTGATATGACAGTAGCTCCCTTACGMAAGATTCTATGAGAAATGACGCTCC 5711
DB ATGCCGCTGATATGACAGTAGCTCCCTTACGMAAGATTCTATGAGAAATGACGCTCC 893
QY 5712 CGGAAATATTGATGATCCCATGATATAGAGCTTTTAAAGGAAACCCCACTTCACTGC 5771
DB CGGAAATATTGATGATCCCATGATATAGAGCTTTTAAAGGAAACCCCACTTCACTGC 953
QY 5772 CCAACCCATATGCCCCGCAACCTGCTATCTACTTGCACCTCTTTCATGATGCAATAC 5831
DB CCAACCCATATGCCCCGCAACCTGCTATCTACTTGCACCTCTTTCATGATGCAATAC 1013
QY 5832 TCATATTGACAGGAAATTAATCCATGTTCCGAGAGAGAGAGAGAGAGAGAGAGAG 5891
DB TCATATTGACAGGAAATTAATCCATGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1073
QY 5892 CTGTTGACTTACTTCACTCCCAACCTGATATGATGAGAGAGAGAGAGAGAGAGAGAG 5951
DB CTGTTGACTTACTTCACTCCCAACCTGATATGATGAGAGAGAGAGAGAGAGAGAGAG 1133
QY 5952 AAGGAAATCATGTAAGAAAGATATCTCCCACTGACGAGGATGACATGACCTCTAG 6011
DB AAGGAAATCATGTAAGAAAGATATCTCCCACTGACGAGGATGACATGACCTCTAG 1134
QY 6012 CCCCCAAGAGACTAGATCTCTCAAAATCATGAAACCTCCGATACCTACTCGGCT 6071
DB CCCCCAAGAGACTAGATCTCTCAAAATCATGAAACCTCCGATACCTACTCGGCT 1194
QY 6072 GGTAAAGCTATTTTAAATGCAACCTCATGAGGCTCATGAGGCTCGGAGGCAAAACCTAC 6131
DB GGTAAAGCTATTTTAAATGCAACCTCATGAGGCTCATGAGGCTCGGAGGCAAAACCTAC 1254
QY 6132 TAACTGTTGATATGCTCCCTGAACTTCAGCCATATGTTCAATCCCTGATCTGA 6191
DB TAACTGTTGATATGCTCCCTGAACTTCAGCCATATGTTCAATCCCTGATCTGA 1314
QY 6192 ACAATGGAACAATCTTCAAGACAGAAATTAACAACAATCCCTCTTTTAAAGAGAGAGAG 6251
DB ACAATGGAACAATCTTCAAGACAGAAATTAACAACAATCCCTCTTTTAAAGAGAGAGAG 1374
QY 6252 TTCCAAATGGAATTAATCCCACTGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6311
DB TTCCAAATGGAATTAATCCCACTGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434
QY 6312 ATACACAACAATCTCCCAATGATCATGAGTGGGTAATCTCTCCACACAATATGCTGCT 6371
DB ATACACAACAATCTCCCAATGATCATGAGTGGGTAATCTCTCCACACAATATGCTGCT 1494
QY 6372 ACCCTGAGAAATTTTGT 6431
DB ACCCTGAGAAATTTTGT 1554
QY 6432 AGAATCATATGCT 6491
DB AGAATCATATGCT 1614
QY 6492 TTTATACAGTTATGATATTAAGCCCGGCAACAAAGATACCATCTCTCTCTCTCTCT 6551
DB TTTATACAGTTATGATATTAAGCCCGGCAACAAAGATACCATCTCTCTCTCTCTCT 1674
QY 6552 TATAGAGACGAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 6611
DB TATAGAGACGAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 1734
QY 6612 TCAATTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGGATC 6671
DB TCAATTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGGATC 1794
QY 6672 CTTGATCACTTGGCAAGATCACTTAATCTCCCTAGACAGATGCTCTTCAATTCGAG 6731
DB CTTGATCACTTGGCAAGATCACTTAATCTCCCTAGACAGATGCTCTTCAATTCGAG 1854
QY 6732 AGCTTTAGACTTGTCTAACCCGCTGARAAGGGGGAACCTGTTATTTTAAAGGAAAGATG 6791

DB 1914 AGCTTTAGACTTGTCTAACCCGCTGARAAGGGGGAACCTGTTATTTTAAAGGAAAGATG 1973
QY 6792 CTGTTATTTATGTTATCAATCCGGAATCTGCTAGTGAATGTTAAAGAAATTCAGATC 6851
DB CTGTTATTTATGTTATCAATCCGGAATCTGCTAGTGAATGTTAAAGAAATTCAGATC 2033
QY 6852 AATACAACTGATGAG 6911
DB AATACAACTGATGAG 2034
QY 6912 GATGCCCTGATTTCTCCCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6971
DB GATGCCCTGATTTCTCCCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2094
QY 6972 TGAACCTGATTTTAACTCTCTGTTAACTTTGTTCTTCCAGATCAGACTGTTAA 7031
DB TGAACCTGATTTTAACTCTCTGTTAACTTTGTTCTTCCAGATCAGACTGTTAA 2154
QY 7032 ACTACAAATGAG 7091
DB ACTACAAATGAG 2214
QY 7092 GCTGTAG 7151
DB GCTGTAG 2274
QY 7152 AGCTGACAACTCTTCACTGATGTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 7211
DB AGCTGACAACTCTTCACTGATGTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2334
QY 7212 AACCTCCCAAG 7271
DB AACCTCCCAAG 2394
QY 7272 CTGGAATTTCTGAGGATGATTAAGATCTTAAGCTGATGAGAGAGAGAGAGAGAGAGAG 7331
DB CTGGAATTTCTGAGGATGATTAAGATCTTAAGCTGATGAGAGAGAGAGAGAGAGAGAG 2454
QY 7332 CTTTAAACAGGGGCTTGCATTAAGTCAACCTGACCAATCAGAGAGAGAGAGAGAGAG 7391
DB CTTTAAACAGGGGCTTGCATTAAGTCAACCTGACCAATCAGAGAGAGAGAGAGAGAGAG 2514
QY 7392 GCTAATTAAGCAAG 7451
DB GCTAATTAAGCAAG 2574
QY 7452 GAG 7511
DB GAG 2634
QY 7512 TTTGGGCTCCCTCCCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7571
DB TTTGGGCTCCCTCCCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2694
QY 7572 TTGCACTGC 7581
DB TTGCACTGC 2763

RESULT 9
US-09-573-080A-26
Sequence 26, Application US/09573080A
Patent No. 6828097
GENERAL INFORMATION:
APPLICANT: JOAN, KNOLL
APPLICANT: ROGAN, PETER
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENER
FILE REFERENCE: 30307
CURRENT APPLICATION NUMBER: US/09/573, 080A
CURRENT FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 479
SOFTWARE: PatentIn version 3.0


```

; SEQ ID NO 26
; LENGTH: 8399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(839)
; OTHER INFORMATION: hevt9
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Malichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE: 1996-01-26
; US-09-573-080A-26

```

```

Query Match 28.6%; Score 2168; DB 4; Length 8399;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;

```

```

Qy 605 TGATGGGAAAGCTTCCCGCAAGCAAAAACGCCCTAAGACGTATCTGGABAATTGGG 664
Db 1276 TGATAGGAAACACTCAGGACATCAACACGCTCAGCTTGAATGTATCTTAAGCCATTGGG 1335
Qy 665 AMCAATTTGACCTCTCAGACCTAAGAAAGAAAGCACTTATTTCTTCTGCAAGTCCGCT 724
Db 1336 ACCAATTTGACCCCAAAACCTGAAAAGAGGCACTATTTTCTGCACTACAGCCT 1395
Qy 725 GGCACCT-----CTGAGGAAAGTAAATTATA 752
Db 1396 GACCTTAATATCTCTCTGATGGGAAAAATGGCCACTGAGGAAAGTAAATTATCA 1455
Qy 753 ACACCATCTTACAGCTAGACTCTTTGT---AGAAAAGCAAAATGAGTAAGTCCAT 809
Db 1456 ATACTATCTGCAAGTTGACCTTTCTCTTAAGAGGAAAGTAAAGAGTAAATACAT 1515
Qy 810 AAG---TACAACTTTCTTTTCAATTAAGACAACTCAATTAATGTAAAGTGTAT 865
Db 1516 AAGGTATTAACAAGCTTTCTTTTCACTGAAGAGAAATACAACTATGCAAAAGCTTGCAAT 1575
Qy 866 TTATGCCCTAAGAAAGCTTCAAGTCTACCTCCCTATCCAGATCCCGAATCCTTC 925
Db 1576 TTACATCCCAAGAGGAACTTTCACTTACCTCCCAATCTTACGCTCTCTAGAGCTCTC 1635
Qy 926 C--CCAAATTAAGAACCCCTTCAACCCAAATGTCCAAAAAGAGATAGACAAAG 983
Db 1636 CTTCTATTAATGCGAAGCTCTCTCCATCTCCCTGCCAGAGAAATTAAGCAAGAA 1695
Qy 984 GTAACAGTGAACCAAGAGTCCCAATATCCCAATTAAGACCCCTCCCAAGAGTGG 1043
Db 1696 ATCTCAAGAGACCAAAAACCCCTGAGCTATCGGTATGTCCCTT-CAAGCTGTAGG 1754
Qy 1044 AGGAAGAAATTTGGCCCAAGCAGAGTCAATGTGTTTTTCTCCCAAGCTTAAAGCA 1103
Db 1755 GGGTGGAAATTTGGCCCAAGCAGATGATGCTCCCTCTGCTGTGATTTAAAGCA 1814
Qy 1104 AATTAACACAGCTTAAGTAAATCTCAGATAATCTGATGAGCTATATGTGTTTACA 1163
Db 1815 GATCAAGGACAGCTGGGGAAAGTTTCAAGATATCTGATAGGACATAGATGCTTCA 1874
Qy 1164 AGGGTTAGACAATTTCTTGATCTGACATGAGAGATATATATGTCACTGCTAAATCGA 1223
Db 1875 GGGCTAGGGCAAACTTCGATCTCACTTGAAGAGATGT-CATGCTATTTAGATCAAA 1933
Qy 1224 CACTTAACCCCAATGAGAAAGTGCACATTAATCTGAGCTGAGAGTGTGGCATCTCT 1283
Db 1934 CCGTGGCTTTTAATGAAAAAGATGTGCTGTAGCTGACGCTGAGAGTTTGAAGATACT 1993

```

```

Qy 1284 GGTATCTCAGTCAAGTCAATGATATANGATGACAAAGAAAGAAAGAAATGATTCCTCA 1343
Db 1994 GGTATCTTAATGCAAGTAAATG---ATGAATGACAGCCGAAAGAAAGATAAATTCCTCA 2050
Qy 1344 CAGCCAGCARGAGAGTTCCCACTTASACCTCTATGGGACACAGAAATCAGTAACATG 1403
Db 2051 CTGGCCAGCAAGCATCCCAAGTATGATCCCACTGGGACCTC---AATCAGATCAT 2106
Qy 1404 GGAGATTGGGCTGCAGACATTTGCTAATCTGTGTGCTASAAAGACTTAAGAAAACATASG 1463
Db 2107 GGGACCTGAGGTCTTAACATCTGTGACCTGTGTGGAAGGACTTAAGGAAATTTAGG 2166
Qy 1464 AAGAAATCTAAGATTAATCAATGATGTCCACCATTAACAAGGAGAAAGGAGAAATC 1523
Db 2167 AA-AAAGCCATGATTTATTCATGATGATCCACATTAATCA-AGAAAGGAAAGAAATC 2224
Qy 1524 CTACTGCTTTCTGAGAGACTTAAGGAGCATTTGAGAAAGCTGCTCTGTCACTG 1583
Db 2225 CTTCTGCTTTCTTCAAGTGGCTTCAAGAGACCTTAAGAAATATATCTCACTGCACCTG 2284
Qy 1584 ACTCTTGAAGGCAACTAATCTTAAGGCTAAGGCTTATCACTCAGTCACTGCAGACA 1643
Db 2285 AATCACTGAGGGTCAATTAATCTTAAGATTAAGTTATTAATCCCAATCAGACAGATTA 2344
Qy 1644 TTAG-AAAAAATTCAAAAGTCTGCCCTAGAGCCCGAGCAAAACTTAAGAAACCTATTGA 1702
Db 2345 TCAGTAAGAAAGTCCAAAGAACAGCCCTGGGCCCTGAACAAATATGAGGCAATTAATTA 2404
Qy 1703 ACTTGGCACTTGGGTTTTTATTAATGAGATCAGAGAGAGCAGAGCCGAAACAGACAAC 1762
Db 2405 ACCTGGCAACCGGTGTCTTATTAATGAGACCAAGAGGAAACAGGCCAA---AGGAAA 2460
Qy 1763 GGGATTAATAAAAGGCAACCGCTTATGATGATGACCTCAGGCAAGTGAATTTGAGGC 1822
Db 2461 AGAGATCAAGAAAAGCCGAGCAGCTTATGATGAGCCCTCAGACCAACAACTTTGAGGT 2520
Qy 1823 TCTGAGAAAGGAAAGCTGGCAATTAAGATGCTAATAGGCTTCTTCACTGCGGT 1882
Db 2521 TCAGAGAGAGCAGAAAATGAGAGCGCAATCAAGCTGTAGGCTTGTATCAATGTGCT 2580
Qy 1883 CTACAGAGACATTTTAAAGAAATGCTCCAAATGAGTAAGCCGCTTCCGTCATGC 1942
Db 2581 TTACAAAGAACATTTAAAGAAATGCTCCAGTGAAGAAACAGGCTCCCTTATCTGTGT 2640
Qy 1943 CCTTATTTCAAGGAATCACTGGAAGGCCCACTGCCCCAGGAGCAAAAGTCTTTGAG 2002
Db 2641 CTACTATGCCGAGGCAATCACTGGAAGTGCATGCCCCAGAGGACAAAGTTCTAGG 2700
Qy 2003 TCAGAAAGCACTAACAGATGATCCAGACAGAGACTGAGAGGTGCTGGGGCAACGCCCA 2062
Db 2701 TTGGAAGCCCCCAACCTGATGATCCAAACAAGGACTGAAGGTTGCCAGGCAACGCCCA 2760
Qy 2063 TCCATATCCATCAACCTCAAGAGCCCTGGGTATGCTGACCATTTGAGGGCAGAGAA 2119
Db 2761 GCTCATGTCATCAACCTCAAGAGCCCTGAGGCTGATTAACATTAAGAGGCGAGAGAAAT 2820
Qy 2120 GGTGTCTCTGGAACACTGAGTGGGCTTCTTGAATCTTACTCTTCTGCTCCGCAACTG 2179
Db 2821 GACTTCTCTTGAACACTGAGCAAGTCTTCTCAGTGTATATCTCTTCTCGATGATCTG 2880
Qy 2180 TCTCTCAGATCTGTCACTATTTGAGAGGGGTCCNTAAGACGGGCACTCATGATCTTT 2239
Db 2881 TCTCTCAAGGTCTGTATCAATCC---GAGGAATCTGGGGGAGGCTGTATCAAGGT-CTTT 2296
Qy 2240 TTTCCAGCCCACTAAGTATTAACCTGGGAGCTTATCTTTTCAACATGCTTTTCAATTA 2299
Db 2937 CTCCCACTCTCAAGTGT-ATTTGGAGACTTGTCTTTCTGTATGCTTTCTTTGTTA 2995
Qy 2300 TGCTTGAAGCCCACTAAGCTTTGTTAGGAGAGACATTTAGCAAAAAGAGGCGCCATTA 2359
Db 2996 TGCTTGAAGTCCCAACCTTTATTAAGGAGATATAT-TAGCCAAAGCTGAGACTATTA 3053

```

Oy	2360	TACACTGAACTATGGAGAGGAAACACCCGTTTGTGTGTCCTTGAAGAAAGAAAT	2419
Db	3054	TCATCATGAATATGGGGAACAAGTTATCCCATTTTGTGT - CCTTATCTTGGAGGGAAT	3112
Oy	2420	AATCCTGAAGTCTGGGCGCAAGAAAGGACAATATGAGAGGCAAAAGATCGCTCGT	2479
Db	3113	AACCTGGAAGTCTGGGCAATTGGAAAGCAATTTGGAAAGGCAAAA - AATGCCACCAAGT	3171
Oy	2480	TCAAGTTAAACTAAAGGATCCACTCTCTTCCCTACCAAGGAGTACCCCTCAGACC	2539
Db	3112	CCAAATCAGGGCTAAAAAGCTCCACACATTTCTCTTATCAAGGCAATATCCTTAAAGGC	3231
Oy	2540	CAGGCCCAACAAGATTCCAAAGATTGTTAAGACTTAAAGCCCAAGGCTTAAGTAA	2599
Db	3232	TGAAGCTCTTAAAGGATTACAGGATATTTGTTAAACATTTAAAGCTCAAGGCTTAAGTAA	3291
Oy	2600	ACCAATGCAATACCTCCCTGACGTAAATCCGTATGTGATGTAGAGAGGCAAGAAACCCAGTG	2659
Db	3292	GAAATGCAAGCTCCCTGCAACACCC - - - - - AATTTAGAGTACAAAACCAATG	3344
Oy	2660	GACAGTGAAGGTTAGTCAAGATCTCAGATTATCAATGAGGCGCTGTCTTTATA	2719
Db	3345	GTCATCTGA - GACTGTGCAAGATTTAGACTCATAT - GAGGACGTAATTCCTCATATA	3402
Oy	2720	CCGACCTGACCTAGCCCTTATCTGTGMYTTCCCAATACCAAGAGACAGATGTT	2779
Db	3403	TCGATTTGATCCCAACCCCTATACCTCTTTCTCAATACCAAGAGACAGAAATGTT	3462
Oy	2780	TACASTCCTGGAACCTTAAAGATGCTTCTTCTGCAATCCCTGTAATCCTGATCTCAAT	2839
Db	3463	CATGTTCTGGACCTCAAGAGATGCTTCTGTGTTCC - CTGCACTGTGCTCCAGTT	3521
Oy	2840	CTGTGTTGCTTGAAGTACTTCAACCCCAATCATCACTCACTGACTTATTTAC	2899
Db	3522	TCTGTTGCTTGAAGATCCACAGACCAACGTCCTCAATTAATGATGATGCTTCC	3581
Oy	2900	CCNAGGTTCAAGGATGAGTCCCATCTATTGGCCAGGCAATTAGCCCAAGCTTAGCYA	2959
Db	3582	ACAAAGGTTTAAAGGATAGCCCTCACTGTGTTGTGAGGCACTGGCCCAAGATTAAGGCA	3641
Oy	2960	RTYATCATACCTGACACTTGTCTCTTCRGTAKGTGATGATTTACTTTTGTGTGCT	3019
Db	3642	CTTCTCAAGTCCAGGCACTTGTGCTTCAAGTATGATGATGATTTACTTTTGTGTACAG	3701
Oy	3020	TTCAAAAACCTGTGCGATCAAGGCAACCAAGCCTCTTMAATTTCTCGCATCTGCG	3079
Db	3702	TTCAAGAGCTCATGCGACGACGACGCTACTTATGATCTTTGAATTTCTTACTTAATCAAG	3761
Oy	3080	CTACAGGTTTCAAAASARAGCTCATCTGCTCAGAGCAGGTTAATACTTAGRCT	3139
Db	3762	GTACAAAGCATCTAAGTCAAGTGAAGTGTGCTTACAGAGGCTAATAATCTAAGGCT	3821
Oy	3140	AATATTATCAAAAGCACARGGCCCTCAGTGAAGAAATATCCAGCTATATCTGCTTA	3199
Db	3822	AATCTTAGCCAAAGGACCAAGGCGCTCAGCAAGAAATGAAATACGCTTAATCTGCTTA	3881
Oy	3200	TCTCATCTCAAAACCTTAAAGCACTAAGBRRTTCTGTGCTTAATTAAGGTTCTGCG	3259
Db	3882	TCTTTGCCCTTAAGACTTAAACAGTTGACGGGGTTCCTTGAAATCAACGGGCTTTTGCG	3941
Oy	3260	AAATATGATTCCCAAGGTTGACRAAATAGCCAGGCTATTAAATCAATAATTAAGAA	3319
Db	3942	ACTATGGA - TCCCTGATATCAATGATATGTAGGCGCCCTCATATCTTAATCAAGAGGA	4000
Oy	3320	CTCAGAAACCAATACCAATTTATTAAGTGAATACGAAATGAAATGATGCTTTCCAGG	3379
Db	4001	CCAGAGGCGCAATATCTTATCTAGTGAATGTATACAGGGCGCAAAACAGCTTCAAAA	4060
Oy	3380	CCCTTAAGAAGCCTTAACCCAGTCCCAAGTGTATGATGTTGCAADRGGCAAGACTT	3439
Db	4061	CCTTAAAGCAGGCGCTGTA - - - CAAAGTTCACCTTTAAGCCTTCCAC - - - - - AATCTT	4111
Oy	3440	TTSTTATATATCACGAAAAAACGAAATAGCTCTGAGATCTCTTACAGAGTCCAG	3499

Db	4112	CTTTTAAATGTCACG-AAGAGCAGGGATAGCTTTGAGTCTTAACTCAGACTCTTG	4170
Qy	3500	GGATGAGCTTGCAACCCYRTGGCCRYACCTGASTAAGGAAATGATGTAGTGGCAAGGGTT	3559
Db	4171	GGACAACCCCAACACAGTGGCATTAAGTAAAGGAAATGTGATGTAGTGCAGAAAGGCT	4230
Qy	3560	GRCYTCATTTGTTAATYGGGTAGTGTGGCACTAGCACTYTCATGTACGTAGCAAGCTTAA	3619
Db	4231	GGCTCACTTTTAAAGGTATGGTTCAGCACTGGCCATCTTCATGTCAGTCAAGGCTATCA	4290
Qy	3620	TAAATCAGGGAGAGATCTTACTGTGTGCACTCTCATGTAGKGTGAATGCGATCACTG	3679
Db	4291	TAAATCAGGAA-AGATCTCATCTGTCTGACTCATGTGTAAATGGCATACTAGGTG	4349
Qy	3680	CTAAAGGAGACTTGTGGCTGTCAACAATGTTACTTAAATTCAGGCTTATTACTTG	3739
Db	4350	CCAAAGGAATTTATGCTATGCACTAACCACTACTTGTGATTCAGGCACTACTCTTG	4409
Qy	3740	AARGGCCAGTGTGCACTGTGCACTTTGGCACTCTTAAACCAAGTCAATTTCTCCAG	3799
Db	4410	AGGACCAAGTCTCAATATCTCAATGTGTGGCCCTCAACCTTGCACTTTCTCCAG	4469
Qy	3800	ACATTAAGAAAGATATBAATATACGTGTCAAAATTAATTTCTCAAACCTATGCACTC	3859
Db	4470	AGAAATGGGAACCAATTTAGCATGACTGCCACAAATTAACGTCCAGACTTATGCCACC	4529
Qy	3860	GAGGGGACCTTGTAGAGTTCYTTGACTGATCCYAGCTTCACTTTGATCTGATGCA	3919
Db	4530	AAGATGATCTCTTAGAGTCCCTTAGCTAATCTGACCTT-AACTATATACCGATGGA	4588
Qy	3920	AGTTCCTTTGTAGAAAAGACCTTCGAAAAGYGGGGATGCAAGTGTCACTGTATATGGA	3979
Db	4589	AGTTCATTTGAGAAAATGGGAATATGAAGGCAAGTTGTGACATAGTATGTGATGTA	4648
Qy	3980	ATATTTAAAGTATCCCTCACTCCAGGAACATAGTCTATGCTGCGCAACTAATAGCC	4039
Db	4649	GTACCTGAAGTAAAGCTCTTCCCCAGGGACAGCACCAAGTTAACAGAACTAGTGGCA	4708
Qy	4040	YTCAATYKGGCACTAGAAATTAGAGNAGRAAAAAGGYAATAATATATACAGACTTAR	4099
Db	4709	CTTACCTGAGCTTAGAACTGGGAAAAGGAAAAGAAATATATGTGATACAGATAGCAG	4768
Qy	4100	TATGCTTACTAGTCTCATGCTCCATGTCATGTCATATGSRABAAAAGGAAATCTTA	4159
Db	4769	TATGCTTATCTAATCCCTCACTGCCCCATGCTGCATATGAAAAGAAAGGAGTCTTA	4828
Qy	4160	TCYAGAGRAACAACCTATCAMAACCTACAGAAAGCCATTGABATTAATATAYGCGTACAG	4219
Db	4839	TCCGGGGTACCCCCATTAATATGCCACAAAGAGTTATGAGTTATTGCACTGCATC	4888
Qy	4220	AAACTABAGAGTGTGMAATCTTCACTGCTGCGGTCTCATMAAAAGGAAAGRAAAGGAA	4279
Db	4889	AAAGCCAGAGAGGTGTGCACTTTTCACTGCAAAAGCATCAGAAAGGTGAAGGAAAG	4948
Qy	4280	ATASAAAGGAATYSCAAGCAKATATTGAAGCMAAAAGAGTGTGCAAGCAGACCTTCA	4339
Db	4949	CGAAGGAACCGTCAAGGAGATGTGAGGCCAAAATTTGCTGCAGGTGGATPCTCCA	5008
Qy	4340	TTAGAAATGCTTATTAATCTTCCCTTAGTATAGGGTATCCCTTCCGGGAAACCAAGCC	4399
Db	5009	TTAGAAATCTTATAGAAAGGACCTTTGAAATGGAACAAACCTTCCAAAGATTAAGCC	5068
Qy	4400	CAGTACTCAGAGGAAACAGAAATGGGGAACCTCAGAGG--CAGTTTCTCCCTCG	4456
Db	5069	CAGTATTCCTCAATATGAACAGAAATGGGACTCTCAATGGGGGCAATGTTTCTCCCTCA	5128
Qy	4457	GAGCGGTAGCCACTGAAGAGGAAAAATCTTTGCTGTGCATATTCATATGAAATTA	4516
Db	5129	GGGTGTTAAGCAAAAAGAGAAAGTACTTATATCCGAAAGCACAGTGGAAAAATA	5188
Qy	4517	CTTAAACCTTCACTAACTTTCACTTGGAGTCGATAGCACCATCAATATGGCCAAA	4576

QY 2608 TAACCTCCGCAATATCCGATGAGATTGAGAGGCAACAGAACCCAGTGCACATGG 2667
 DB 29084 GCAAGTCCCGCAACACCC-----AATTCTAGGAGTACAAAAACGTGACGTG 29136
 QY 2668 AGGGTATGAGCAAGATCTCAGGATTATCAATGAGAGCCGTTGCTTTATATCCGAGCG 2727
 DB 29137 A-GACTAGTGAAGATCTTAGCTCATCAT-GAGGCACTAATCTCATATATCCAGTTG 29194
 QY 2728 TACCTAGCCCTTATCTGTGMYTTCCCAATACAGAGAGAGAGAGTGGTTTACASTCC 2787
 DB 29195 TACCCACCCCTTATACCTGCTCTCTCAATACAGAGAGAGAGATGGTTTCAAGGTC 29254
 QY 2788 TGGACCTTMAAGATGCTTCTTCTGATCCCTGATATCTGATCTGATCTGATCTGATCTG 2847
 DB 29255 TGGACCTTMAAGATGCTTCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 29314
 QY 2848 CCTTGAAGATCTTCAACCCACATCTCACTGACCTGACCTGATCTGATCTGATCTGATCTG 2907
 DB 29315 CCTTGAAGATCTTCAACCCACATCTGACCTGACCTGATCTGATCTGATCTGATCTGATCTG 29374
 QY 2908 TCAGGATAGTCCCATCTATTTGCGCAGGCACTTACCCCAAGCTTGAATCTGATCTGATCTG 2967
 DB 29375 TTAGGATAGTCCCATCTATTTGCGCAGGCACTTACCCCAAGCTTGAATCTGATCTGATCTG 29434
 QY 2968 ACCTGACACTCTTCT 3027
 DB 29435 GTCCAGGCACTCTGCT 29494
 QY 3028 CCTTGGCATCTCAAGCAACCCCAAGCTCTTMAATTTCCGCGACCTGCGCTGATCTGATCTG 3087
 DB 29495 CCTTGGCATCTCAAGCAACCCCAAGCTCTTMAATTTCCGCGACCTGCGCTGATCTGATCTG 29554
 QY 3088 TTTCGA-AACSAARAGCTCACTCTGCTCAACAGAGTTAAATCTTAAAGCTTAAATCTTAA 3146
 DB 29555 TCTCTAGTCTGAAGGCTCCGAGCTTTCCTCAACAGAGTTAAATCTTAAAGCTTAAATCTTAA 29614
 QY 3147 TCCAAAGGCAACGAGGCTCTCAAGTGAAGATCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 3206
 DB 29615 GCGAAAGGCAACGAGGCTCTCAAGTGAAGATCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 29674
 QY 3207 CYCAAAACCTTAAAGCACTAAGRRTTCTGCGRTAAAGSYTTCCGCGAATG 3266
 DB 29675 CTTCAAGCACTTAAAGATGTTGGGGTCTCTGGAATTACTGCTTGGCCCACTATGG 29734
 QY 3267 ATTCCCGAGTWTGGRAAATAGCCAGGYCATTAATTAATTAAGGAACTCAAG 3326
 DB 29735 A-TCCCGGATACAGGAGTACCCAGGCACTCTGAGCTTAATCAAGGACCCAGAG 29793
 QY 3327 AGCCATATCCATTTATATAGATGAYAMCTGAAGTAAAGTGGCTTCCAGGCCCTTA 3386
 DB 29794 GCGAATATC-----TCTGTTGGGAATCAGAGGCAAGACCTTCAAAACCTTAA 29845
 QY 3387 AGAAGCCCTTAAACCCAGTCCAGTGTAAAGTTGCAACRGGGCAAGCTTTSTTA 3446
 DB 29846 GGAAGGCTT-AGTAAAGTCCAGCTTAAAGCTTCCACAGTCAAAAGCTTCTCTTA 29903
 QY 3447 TAVRTCAAGAAAAAAGCAAGAAAGTCTGAGTCTTCAACAGTCCRAGGAGYAG 3506
 DB 29904 TAGTACAG-AAGAGCCGGGATAGCTTCTGAGTCTTCAAGTCTGAGGCAAC 29962
 QY 3507 CTTGCAACCTTGAAGCTGATCTGATGAAGAAATGATGATGAGGCAAGGCTTGTCTCA 3566
 DB 29963 CCAACCAACGATGATCTTAAAGTAAAGAAATGATGATGATGATGATGATGATGATGATG 30022
 QY 3567 TTTTATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3626
 DB 30023 CTGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 30082
 QY 3627 GGGAGAGATCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3686
 DB 30083 AGGAAAGATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 30142

QY 3687 AGACTTGGCTGTCAAGCACTGTTTACTTAATRTCAGGCTTATCTTAAAGGCC 3746
 DB 30143 AAGTTATGGCTATCAGACACCCACTTATGATCAGAGGCTCTCTTGAAGACC 30202
 QY 3747 AGTGTGACACTGTGCACTTGTGCAACTTAAACCAAGTCAATCTTCCAGACATGA 3806
 DB 30203 CATGTTAAATTAAGCAAGTGTGTGGCCCTCAACCCCTGCTTCTCCAGAGATGG 30262
 QY 3807 AGAAAGTAAATTAATCTGTCAACATTAATTTCTGAAACCTTATGCACTGAGGGA 3866
 DB 30263 GGAACCAATTAAGCACTGCTGCAACAAATTAATCTGCACTTATGCTGCGGAGATGA 30322
 QY 3867 CCTTATAGATGCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 3926
 DB 30323 TCTGTAAAGTCCCTTAACTATCTGACCTT-AACTATATCTGATGAGATTCAT 30381
 QY 3927 TTGTAGAAAAAGCACTTGAAGAAAGGCTATGCAAGTGTCACTGATTAATGAAATATTTG 3986
 DB 30382 TTGTGAGAAATGGGATAGCAAGGCGAGTTACGCCATAGTATGATGATGATGATGATGATG 30441
 QY 3987 AAAGTATCCCTCACTCCAGGAATAGTGTGATGATGATGATGATGATGATGATGATG 4035
 DB 30442 AAAGCAAGCTTATCCCAAGACCAAGTCCCAAGTATGATGATGATGATGATGATGATG 30490

RESULT 12
 US-08-979-847B-88
 / Sequence 88, Application US/08979847B
 / Patent No. 6582703
 / GENERAL INFORMATION:
 APPLICANT: PERON, HERVE
 BESEME, FREDERIC
 BEDIN, FREDERIC
 PARANHOS-BACCALA, GLAUCIA
 KOMURIAN-PRADEL, FLORENCE
 JOLIVET-REYNAUD, COLETTE
 MANDRAND, BERNARD
 GARSON, JEREMY
 TIKE, PHILIP
 TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
 ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPH
 THERAPEUTIC PURPOSES
 NUMBER OF SEQUENCES: 210
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OLIF & BERRIDGE, PLC
 STREET: P.O. BOX 19928
 CITY: ALEXANDRIA
 STATE: VA
 COUNTRY: USA
 ZIP: 22320
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/979,847B
 FILING DATE: 26-No. 6582703-1997
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BERRIDGE, WILLIAM P.
 REGISTRATION NUMBER: 30,024
 REFERENCE/DOCKET NUMBER: WPB 39046A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-836-6400
 TELEFAX: 703-836-2787
 INFORMATION FOR SEQ ID NO: 88:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2364 base pairs
 TYPE: nucleotide
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 88:

Db	1125	CAGGGCCCTCAGTGAAGAAAGTATCCAGGCTTACTGGCTTATCTCCATCCAAAACCT	1189
OY	3218	AAAGCAACTAAGGRRTTCTTGGCTAAVAGYTTCTGCCAATAATGATTTCCCAAGT	3277
Db	1185	AAAGCAACTAAGGGGTCTTGGCATTAACAGTTCCTCCAAAACAGATT-CCCAAGT	1243
OY	3278	MTGGCAAAATAGCCAGGCACTTATATACASTAATTAAGAACTCAGAAAGCATATCC	3337
Db	1244	ACAGCCCAATAGCCAAACCTTATATACCTAATTATGAAACTCAGAAAGCAATACCT	1303
OY	3338	ATTATTAAGATGAGAYAMCTGAAYVRAAGTGGCTTCCAGGCCCTTAAGAGGCTTTA	3397
Db	1304	ATTACTAATGATGACACCT--ACAGAAATGGCTTCCAGGCCCTTAAGAGGCTT--	1358
OY	3398	AACCCAAAGTCCCAAGTGTAAAGTTTGGCAACRGGGCAAGACTTTSTYATVATACAGA	3457
Db	1359	AACCCAAAGCCCAAGTGTACAGTTGGCAACAGGGCAAGATTTTCTTATATGCAAG-	1417
OY	3458	AAAAAACAGAAVAGCTGRRGAGTCCTAACACRCCRAGGAYAGCTTGCAACCYR	3517
Db	1418	AAAAAACAGAAATAGCTTAGAGATCTTAACCAAGGTCTCAGGGATGAGCTTGCAACCG	1477
OY	3518	TGGCAVACCTGASTAAGAAATGATGTAGTGGCAAAAGGTTGRCYTCATTGTTTAYGG	3577
Db	1478	TGGTATACCTGAGTAAGAAATGTATGTAGTGGCAAAAGGTTGGCTCATGTTTATAGG	1537
OY	3578	TAGTGGTGGCAGTACAGAGTYTATGTATCTGAAGCAATTAAATATACAGGAGAGATC	3637
Db	1538	TATATGAGGCACTACAGTCTAAGATATCGAAGCAGTTAAATATATACAGGAAGAGATC	1597
OY	3638	TTACTGTGGGACATCTCACTGAKGGAAVYRGATCTACCTGCTAAAGAACTTGTGGC	3697
Db	1588	TTTCTGTGGGACATCTCACTGATGGAACGGATVCTCAGCTAAAGAACTTGTGGT	1657
OY	3698	TGTCAGACAACGTTTACTTAATATATACAGGCTCTATTACTTGAAGGCGAGTGGCAC	3757
Db	1658	TGTCAGACAACCATTTACTTANTATACAGGCTCTATTACTTGAAGGCAAGCTGTGNAC	1717
OY	3758	TGTGCACTTGTGCAACTCTTAAACCAAGYCNCAATTTCTTCCAGACAATGAAAAAGATR	3817
Db	1718	TGGGCACTTGTGCAACTCTTAAAC--	1742
OY	3818	AAATATACTGTCAACAARTAAATTTTCCAAACCTATGCCATCTCGAGGGCACTTTTAGRG	3877
Db	1743	-----CAACTTAAGCTGCCAGAAAGATCTTNTAGAG	1776
OY	3878	TTCCYTTGACTGATCCYGCCTTCAACT--TGTAATCGATGGAAAGTCTTTGTGAA	3934
Db	1777	GTCCTTATAGCCAACTTCACTCAACTATATATATCTGATGAAAGTTCGTTGTGAA	1836
OY	3935	AAAGCACTTGGAAAAGGGGGATGCT--AGTGTCACTGATTAATGAAATATTTGAAGT	3991
Db	1837	AAGGATTAACAAAGGAGNAGATATCCATAGGTGTATGATTAAGCAGTACTTGAAGT	1896
OY	3992	AATCCCTT---CACTCCAGGAATAGTGTCTAAGCTGCAAGAACTATATGCTT--CAAT	4045
Db	1897	AAGCTCTTCCCCCCCCAGGAAACAGGCCGCCGTTAGCAGACCTAAGTGCACACCCG	1956
OY	4046	KGGGCACTAGAAATTTGAGAAAGRAAAAAGGGTAATATATATACAGACTCRATATGCT	4105
Db	1957	CGAGCTTATGAACCTTGGAAAAGGAGGATTAATGTGTATACAGATAGCAAGTATGCT	2016
OY	4106	YACCTAGTCNTC-----ATGCCATGMRGCAATATGS	4138
Db	2017	TATCTAATCCGAATATGCCCATGTGTATTATCTAATCCGAATATGCCCATGTGCAATATGG	2076
OY	4139	ARAGAAAGGAATTTCTTAATCTTCTYAGGAAACACTATACMAACATCAGAAAGCAATTAGG	4198
Db	2077	AAAGAAAGGAGTTCTTAACCTCTGGGGGAAACCCCATTAATACCAACAAATTATCATG	2136
OY	4199	ARATTTATAYGGCGGTACAGAAACCTARAGAGGGMAGCTTAACCTGCGGGGCTAT	4258
Db	2137	GAGTTATTTGCACAGTGCATAAAATCTAAGAGGTGGAAGCTTTACATGCTCAAGCAT	2196

QY 4259 CANAAGGAAAGRAAAGGGAATASAGRGAAITGCCAA 4297
 ||||| : ||||| : |||||
 Db 2197 CAGAAAAGGAAAGGGGAGAAGCAGCATTAAGTGGCTA 2235

No + p rear

RESULT 14
 US-09-374-766-57
 : Sequence 57, Application US/09374766
 : Patent No. 6579526
 : GENERAL INFORMATION:
 : APPLICANT: HEYRE PERRON
 : APPLICANT: FREDERIC BESEME
 : APPLICANT: FREDERIC BEDIN
 : APPLICANT: GLAUCIA PARANHOS-BACCALA
 : APPLICANT: FLORENCE KOMURIAN-PRADEL
 : APPLICANT: COLETTE JOLIVERT
 : APPLICANT: BERNARD MANDRAND
 : TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
 : TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
 : TITLE OF INVENTION: THERAPEUTIC PURPOSES
 : NUMBER OF SEQUENCES: 92
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: OLIFF & BERRIDGE
 : STREET: 700 South Washington Street, Suite 300
 : CITY: Alexandria
 : STATE: Virginia
 : COUNTRY: U.S.A.
 : ZIP: 22314
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/374,766
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/691,563
 : FILING DATE: 02-AUG-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Berridge, William P.
 : REGISTRATION NUMBER: 30,024
 : REFERENCE/DOCKET NUMBER: WPB 38588
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 703-836-6400
 : TELEFAX: 703-836-2787
 : INFORMATION FOR SEQ. ID NO: 57:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2391 base pairs
 : TYPE: nucleotide
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : US-09-374-766-57
 :
 : Query Match 18.7%; Score 1415.8; DB 4; Length 2391;
 : Best Local Similarity 78.4%; Pred. No. 0;
 : Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;
 :
 : 2021 ATGATCCAGACGAGCACTGAGGGTCCCTGGGGCAAGCCCATGCCATGCATCAACCCCTC 2080
 : |||||
 : 1 ATGATCCAGACGAGCACTGAGGGTCCCTGGGGCAAGCCCATGCCATGCATCAACCCCTC 60
 :
 : 2081 ACAGAGCCCTGGGTATGCTTGAACATTGAGGGCCAGGAAGT---TGTCTCTGGACACT 2137
 : 61 ACAGAGCCCTGGGTATGCTTGAACATTGAGGGCCAGGAAGTATGCTTCTGGACACT 120
 :
 : 2138 GGTGCGGTCTTCTTACTTCTTACTCTTCTGTCTCCGGACCACTGCTCCAGATCTGCACT 2197
 : 121 GGGGAGGCTTCTCTCACTCTTACTCTTCTCTCTCGGACCACTGCTCTCCAGATCTGCACT 180
 :
 : 2198 ATTCTGAGGGGCTCNTAAGCGGCACTCATAGATCTTCTTCTCCAGCCACTAAGTTA 2257


```
Db 1359 AACCCAGGCCCGAGTGTCCAGCTTGCCACAGGCGAAGATTTTCTTATATGCGCAG- 1417
QY 3458 AAAAAACAGAAAYAGCTCTRGAGTCCCTTACACAGATCCRAGGAAVAGCTTCAACCYR 3517
Db 1418 AAAAAACAGAAATAGCTTCTAGAGTCTTACGCAAGTCTCAGGGAATAGCTTGCACCCG 1477
QY 3518 TGGCRYACCTGASTAAGAAAYTGAATGATGTGGCAAGGGTTGRCYTCATTGTTAYGGG 3577
Db 1478 TGGTATACCTGAGTAAAGAAATGATGTAGTGGCAAGGGTTGGCTCATNGTTATAGG 1537
QY 3578 TAGTGTGCGAGTAGCAGTCTKTAGTATCTGAAGCAGTTAAATATATACAGGAGAGATC 3637
Db 1538 TATAGNAGCGAGTAGCAGTCTNAGTATCTGAAGCAGTTAAATATATACAGGAAAGATC 1597
QY 3638 TTAAGTGTGAGATCTCATAGAGTGAAYRGCATPACTCACGTCTAAAGSAGACTGTGGC 3697
Db 1598 TTNCTGTGAGACATCTCATAGATGTGACGCAATACTCACTGCTAAGGAGACTTGTGT 1657
QY 3698 TGTGACACAACVGTTTACTTAAATRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRAC 3757
Db 1658 TGTGACACAACATTTATCTTAATATCAGGCTCTATTACTTGAAGCCAGTGTGNAGC 1717
QY 3758 TGTGCACTTGTGCACTCTTAACCCAGTCATTTCTTCCAGACAATGAAAAAGATAR 3817
Db 1718 TGGCACTGTGTGCACTCTTAAAC----- 1742
QY 3818 AAYATAAGTGTCAACARTATTTCTCAACCTATGSCACTCGAGGGACCTTATAGRG 3877
Db 1743 -----CAACCTATGTCTGCCAGAAAGATCTTTNTAGAG 1776
QY 3878 TTCCYTTGACTGATCCYAGCTTCACT---TGTATACTGATGGAAGTCTTTGTAGAA 3934
Db 1777 GTCCCTTACCAACCTGACCTCACTATATATATCTGATGGAAGTGTGTGTAGAA 1836
QY 3935 AAAGACTTGAAGAGTGGGTATGC---AGTGTCACTGATATATGGAATATTTGAAGT 3991
Db 1837 AAGGATTTACAAAGGAGATATNCCATAGGTGTGTAGTATAAAGCATCTTGAAAGT 1896
QY 3992 AATCCCT---CACTCCAGGAACATGCTAGCTRGCAGAACTAATAGCCYT---CAYT 4045
Db 1897 AAGCTCTTCCCTCCCGCAGGACCGCCGCCCTTAGCAAGACTAGTGGCACTGACCCG 1956
QY 4046 KGGGCACTAGAAATTAGAGAAAGRAAAAGGVAATAATATATACAGACTTARATAGCT 4105
Db 1957 CGAGCTTAAACTTTGAAAAGGAGAGATTAATGTGTATACAGATAGCAAGTATGCT 2016
QY 4106 YACCTAGTCNTCC-----ATGCCCATGMRGCAATAGS 4138
Db 2017 TATCTAATCCGAATGCCATGTGTTATCTAATCCGAATGCCATGTGCAATAGG 2076
QY 4139 AAGAAAGGGAATCTTAATCTTGYAGRAACACTATCAMAACATCAGAAAGCCATTAGG 4198
Db 2077 AAGAAAGGAGTCTTAACCTCTGGGGAAACCCCATTAATACCAAGTTAATCATG 2136
QY 4199 AATATTATTAYTGGCGWTACAGAAACTARAGAGGTGMAAGTCTTACACTGCGGGGTCAAT 4258
Db 2137 GAGTTATTGACACAGATGCAAAAACCTCAAGAGGTGMAAGTCTTACACTGCGAAAGCCAT 2196
QY 4259 CAAAAAGAAAGRAAAGGAATASAAAGRAAYTGCCAA 4297
Db 2197 CAGAAAAGGAAAGGAGAGAGACGATTAAGTGGCTA 2235
```